

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 08-190560

(43)Date of publication of application : 23.07.1996

(51)Int.Cl.

G06F 17/27

(21)Application number : 07-001950

(71)Applicant : FUJITSU LTD

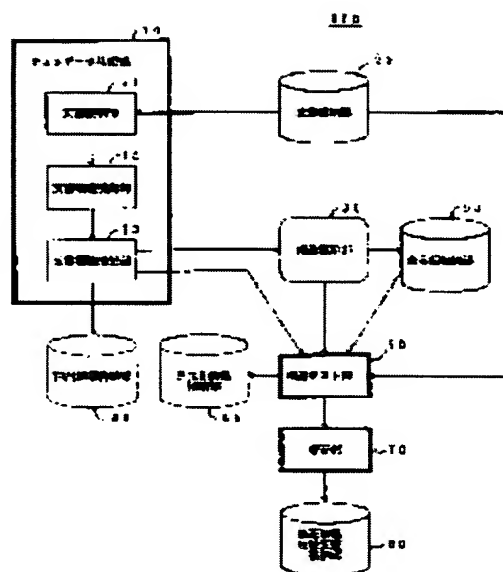
(22)Date of filing : 10.01.1995

(72)Inventor : ATSUMI YASUYO

(54) STRUCTURE TEST DEVICE

(57)Abstract:

PURPOSE: To provide a structure test device for structured document capable of easily processing the structural test of the structured document at a high speed.
CONSTITUTION: This device is provided with a test data generation means 10 which generates test data for display of the structure of the structured document, a test means 50 which uses test data acquired from the test data generation means 10 to extract the error of the structure of the structured document and outputs it as the test result, and a correction means 70 which corrects the error part based on the test result of the test means 50.



LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's

decision of rejection]

[Date of extinction of right]

Copyright (C); 1998,2003 Japan Patent Office

(19)日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開平8-190560

(43)公開日 平成8年(1996)7月23日

(51)Int.Cl. ⁶	識別記号	序内整理番号	F I	技術表示箇所
G 0 6 F 17/27		9288-5L	G 0 6 F 15/ 20	5 5 0 E
		9288-5L		5 5 0 F

審査請求 未請求 請求項の数13 O L (全 26 頁)

(21)出願番号 特願平7-1950

(22)出願日 平成7年(1995)1月10日

(71)出願人 000005223

富士通株式会社

神奈川県川崎市中原区上小田中4丁目1番
1号

(72)発明者 渥美 康代

神奈川県川崎市中原区上小田中1015番地
富士通株式会社内

(74)代理人 弁理士 伊東 忠彦

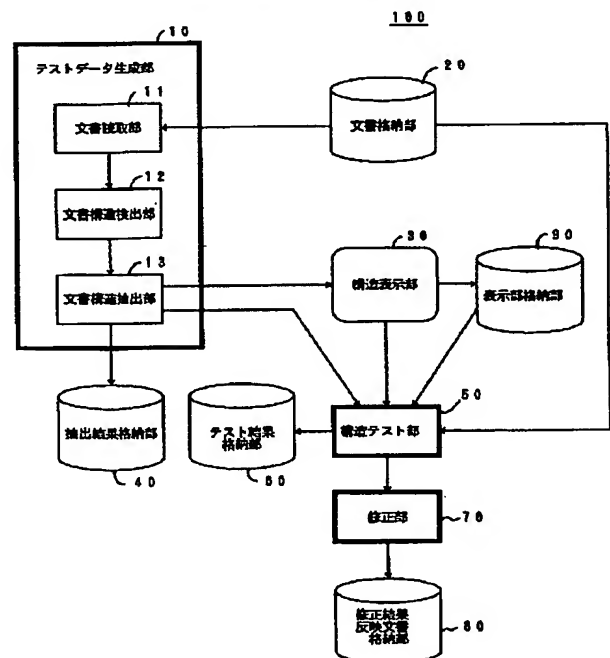
(54)【発明の名称】 構造テスト装置

(57)【要約】

【目的】 本発明の目的は、構造化文書の構造的なテストを簡易で高速に処理することが可能な構造化文書における構造テスト装置を提供することである。

【構成】 構造化文書の構造を表示するためのテストデータを生成するテストデータ生成手段10と、テストデータ生成手段10より取得したテストデータを用いて、構造化文書の構造のエラーを抽出してテスト結果として出力するテスト手段50と、テスト手段50のテスト結果に基づいて、エラー部分を修正する修正手段70とを有する。

本発明の一実施例の構造テスト装置の構成図



1

【特許請求の範囲】

【請求項1】 文書作成の際に構造的な関連により該文書のテストを行う構造テスト装置において、前記文書の構造を解析するためのテストデータを生成するテストデータ生成手段と、前記テストデータ生成手段より取得した前記テストデータ、または、入力された文書を用いて、前記文書の確認すべきまたは、エラーである構造表示データ及び導入部を抽出してテスト結果として出力するテスト手段と、前記テスト手段の前記テスト結果に基づいて、前記エラー部分を修正する修正手段とを有することを特徴とする構造テスト装置。

【請求項2】 前記テストデータ生成手段は、前記文書を読み込む文書入力手段と、入力された前記文書の各構成要素の各文書構成上の該構成要素の内容を検出する構造検出手段と、前記構造検出手段で検出された前記構成要素の内容毎に前記入力された文書中の構成表示データ及び導入部分を抽出する構造抽出手段と、前記構造抽出手段により抽出された前記構成要素の内容毎に構成表示データ或いは、該構成表示データ及び前記導入部分を表示する構造表示手段を有する請求項1記載の構造テスト装置。

【請求項3】 前記構造表示手段は、前記構成要素の件数を表示し、構成表示データのみ、または、該構成表示データに導入部分を組合せて表示されている構成要素を選択させるための選択手段を含む請求項2記載の構造テスト装置。

【請求項4】 前記構造表示手段に表示した内容に加えて、文書の構造情報を併せて保持する記憶手段を有する請求項2記載の構造テスト装置。

【請求項5】 前記テスト手段は、文書全体または、テストを希望する箇所が指定されると、構成表示データ間の構成要素の関係をチェックする見出しチェック手段と、前記導入部の主語が同じ構成要素内で同じ形式を有しているか、及び文末が統一されているかをチェックする導入部チェック手段と、前記見出しチェック手段及び前記導入部チェック手段の結果を出力するテスト結果出力手段を有する請求項1記載の構造テスト装置。

【請求項6】 前記見出しチェック手段は、上位の構成表示データの項番と内容をテストする項番・内容テスト手段と、前記テスト手段において、前記構成表示データの内容と該内容が属する構成要素に対応する所定のルールを照合し、所定のルールと一致する場合には、確認する必要がある構成要素として格納する第1の格納手段と、前記所定のルールと不一致の場合には、照合する前記所定のルールの構成要素を下位または上位に変化させて照

2

合し、どの構成要素も一致しない場合には、エラーと見做す構成要素として格納する第2の格納手段を有する請求項5記載の構造テスト装置。

【請求項7】 前記項番・内容テスト手段は、各項番の構成要素がシーケンシャルにインクリメントされているかをチェックする項番チェック手段を含む請求項6記載の構造テスト装置。

【請求項8】 前記導入部チェック手段は、前記導入部間においても行う請求項6記載の構造テスト装置。

【請求項9】 前記導入部チェック手段は、文書内の全ての前記導入部の文末の表現が常体か敬体であるかを個数から判定し、全体の文書の導入部の文末の形態を判定する手段と、全体の文書の導入部の文末の形態から各々の導入部の文末の形態の正しさを判定し、エラーを指摘する手段とを有する請求項5記載の構造テスト装置。

【請求項10】 前記導入部チェック手段は、導入部の先頭の一文の主語を取り出す主語取出手段と、該主語取出手段より取り出された主語を単語に分解する単語分解手段と、分解された単語が2つ以上であれば、先頭と最後の単語を取り出し、1つであれば、当該単語を取り出す単語抽出手段と、前記単語抽出手段において、抽出された単語の形態を比較して少ない単語の形態をエラーと見做す主語チェック手段を有する請求項5記載の構造テスト装置。

【請求項11】 前記テスト結果出力手段は、前記テスト結果を表示する手段と、前記テスト結果を保持する手段とを含む請求項5記載の構造テスト装置。

【請求項12】 前記テスト結果出力手段は、前記テスト結果に前記文書内のエラーの位置を含むよう編集する手段を有する請求項8記載の構造テスト装置。

【請求項13】 前記修正手段は、前記テスト結果を表示し、指定された前記文書内の前記テスト結果に示されるエラー位置を検索する位置検索手段と、前記位置検索手段により検索された位置の前記文書内における概念位置のエラー、前記導入部の先頭の一文の主語の誤り及び導入部の先頭の一文の主語を修正する手段及び前記導入部の文末を修正する手段を含む請求項1記載の構造テスト装置。

【発明の詳細な説明】

【0001】

【産業上の利用分野】 本発明は、構造テスト装置に係り、特に、マニュアル等の文書を作成する際に、文書の「部」、「章」、「節」、…等の構成要素の配列が正しいかを判定し、構成要素の配列に誤りがある場合には、必要に応じて修正するための構造テスト装置に関する。

3

【0002】詳しくは、部、章、節、項等の構成要素を有する構造化文書や段落を含む一般の文書において、その順序や階層構造が正しく配置されているか、または、各構成の導入部は、それぞれの構成要素と合致した書き方になっているか等をテストするための構造テスト装置に関する。

【0003】

【従来の技術】従来、文書の構造をテストするための具体的ツールは存在していないため、人間が目視により作成された文書をチェックする方法が採られている。例えば、マニュアルをチェックする際には、最初のページから最後のページまで通して読み、目についた点に対してチェックを行うのが一般的である。

【0004】また、文書を格納している記憶部よりチェック対象の文書を画面上に読出し、さらに、目次構成等の構造的な文章作成方法に合わせ、“部分と部分”及び“全体と部分”の関係でテストを行う。例えば、マニュアルの作成技術に違反しているか否かをチェックする際には、構造的なテストを以下の点においてチェックする。

- ・情報が章、節及びパラグラフのレベルで書き方に一貫性があるか。
- ・標準的な説明形式に合わせて説明が標準化されているか。

【0005】上記の“部分と部分”の関係のテストは、説明の一貫性を保つためのテストであり、

- ・目次の一貫性
- ・導入部の一貫性
- ・文章構成の一貫性
- ・パラグラフ内の一貫性
- ・用語一覧表の表記の一貫性

のうち、目次の一貫性については、以下のような手順でチェックを行う。

1章の目次を読み、2章の目次を読み、…、n章の目次を読む。

1. 1. 1. 1. 2…節の目次を読み、2. 1. 2. 2…節の目次を読み、n. 1. n. 2…節の目次を読む。

1. 1. 1. 1. 1. 2…の目次を読み、1. 2. 1. 1. 2. 2…の目次を読む。

【0006】上記の手順に従って目次を読むことにより、同一の属性を持つ主題に対して同じ形式で書かれているかをチェックする。この手順に従うことにより、目次項目の一貫性をチェックすることが可能である。次に、導入部の一貫性のテストは、以下のような手順でチェックを行う。

1章の導入部の文を読み、2章の導入部の文章を読み、…n章の導入部の文を読む。

1. 1. 1. 2…節の導入部の文を読み、2. 1. 2. 2…節の導入部の文を読み、n. 1. n. 2…節の導入部の文を読む。

4

1. 1. 1. 1. 1. 2…の本文の先頭を読み、1. 2. 1. 1. 2. 2. …の本文の先頭を読む。

【0007】上記のように目次レベルに対応した導入部の文を読むことにより、目次レベルで並列の関係にある導入部の書き方について一貫性をチェックする。次に、文章の一貫性のテストは、同一形式を持つ文章表現、例えば、コマンドの説明部分に、一貫性があるかをチェックする。各説明の部分の形式は、全体的に統一されていなければならない。

10 【0008】パラグラフ内の一貫性のテストは、パラグラフに対しては、パラグラフ技術に合わせて一貫性があるかをチェックする。説明のパターンが同種の情報に対して一貫性があるかをチェックする。用語・図表の表記の一貫性のテストは、用語や図表等の表記に関しては、マニュアル上の記載を見て統一性があるかをチェックする。

【0009】

【発明が解決しようとする課題】しかしながら、上記の従来の方法は、文書のテストを行う場合において、作業を支援するような簡易なツールはなく、一貫性や統一性を確認するには、人間が目視によりチェックしているため、テスト漏れが発生するという問題がある。

【0010】また、目次等を表示して、その目次と構成要素の配列が一致しているかを見る方法において、足りない構成要素や、従属関係については目次等を参照することにより、チェックできるが、各文書の構成要素毎のチェックはできないという問題がある。例えば、見出し部に続く第1段落の主語が正確に使用されているか、各構成要素の文末が統一されているか等のチェックは、マニ
30 ニュアル全体を用語や文の意味を意識しながら行われな
い限り不可能であり、多大の時間を要する。また、模範
的な構造的な文章作成方法に合わせてチェックする際に、
チェックを行う者が全てのチェック項目を各々意識して
全文書を通してチェックできるとは限らない。

【0011】本発明は、上記の点に鑑みなされたもので、人手に依らず文書の構造的なテストを電子的に容易で、しかも高速に処理することが可能な構造テスト装置を提供することを目的とする。また、本発明の更なる目的は、テストを行う際に、文書をテストできる状態のデ
40 ータを取得できる構造テスト装置を提供することである。

【0012】また、本発明の更なる目的は、テスト結果に基づいてエラーを自動的に修正することが可能な構造化テスト装置を提供することである。

【0013】

【課題を解決するための手段】本発明は、文書作成の際に構造的な関連により該文書のテストを行う構造テスト装置において、文書の構造を解析するためのテストデータ
50 を生成するテストデータ生成手段と、テストデータ生成手段より取得したテストデータ、または、入力された

文書を用いて、文書の確認すべきまたは、エラーである構造表示データ及び導入部を抽出してテスト結果として出力するテスト手段と、テスト手段のテスト結果に基づいて、エラー部分を修正する修正手段とを有する。

【0014】また、上記のテストデータ生成手段は、文書を読み込む文書入力手段と、入力された文書の各構成要素の各文書構成上の構成要素の内容を検出する構造検出手段と、構造検出手段で検出された構成要素の内容毎に入力された文書中の構造表示データ及び導入部分を抽出する構造抽出手段と、構造抽出手段により抽出された構成要素の内容毎に構造表示データ或いは、構造表示データ及び導入部分を表示する構造表示手段を有する。

【0015】また、上記の構造表示手段は、構成要素毎の件数を表示し、構造表示データのみ、または、構造表示データに導入部分を組合わせて表示されている構成要素を選択させるための選択手段を含む。また、上記の構造表示手段に表示した内容に加えて、文書の構造情報を併せて保持する記憶手段を有する。

【0016】また、上記のテスト手段は、文書全体または、テストを希望する箇所が指定されると、構造表示データ間の構成要素間の関係をチェックする見出しチェック手段と、導入部の最初の一文の主語が同じ構成要素内で同じ形態を有しているか、及び導入部の文末の表現が統一されているかをチェックする導入部チェック手段と、見出しチェック手段及び導入部チェック手段の結果を出力するテスト結果出力手段を有する。

【0017】また、見出しチェック手段は、構造表示データの項番と内容をテストする項番・内容テスト手段と、項番・内容テスト手段において、構造表示データの内容とその内容が属する構成要素に対応する所定のルールを照合し、構造表示データが所定のルールと一致する場合には、確認する必要がある構成要素として、格納する第1の格納手段と、所定のルールと不一致の場合には、照合する所定のルールの構成要素を下位または上位に変化させて照合し、どの構成要素も一致しない場合にはエラーと見做す構成要素として格納する第2の格納手段を有する。また、項番・内容テスト手段は項番のチェックを行う際に、各項番の構成要素がシーケンシャルにインクリメントされているかをチェックする項番チェック手段を含む。

【0018】また、導入部チェック手段は、文書内の全ての導入部の文末の表現が常体である個数、敬体である個数をカウントし、それを文書全体の導入部の文末の形態を判断する材料とする。また、導入部チェック手段は、導入部の先頭の一文の主語を取り出す主語取出手段と、主語取出手段により取り出された主語を単語に分解する単語分析手段と、分解された単語が2以上であれば、先頭と最後の単語を取り出し、1であれば、当該単語を取り出す単語抽出手段と、単語抽出手段において、抽出された単語の形態を比較して少ない形態をエラーと

見做す主語チェック手段を有する。

【0019】また、上記のテスト手段は、テスト結果を表示する手段と、テスト結果を保持する手段とを含む。また、上記のテスト手段は、テスト結果が、文書内のエラーの位置を含むように編集する手段を有する。

【0020】また、上記の修正手段は、テスト結果を表示し、指定された文書内のテスト結果に示されるエラー位置を検索する位置検索手段と、位置検索手段により検索された位置の文書内における概念位置のエラー、導入部の先頭の一文の主語の誤り及び導入部の文末のエラーを修正する手段を含む。

【0021】

【作用】第1の発明は、テストデータ生成処理、テスト処理及び修正処理を行う文書構造テスト装置の全体に関する発明であり、文書を解析して個々の文書内の部、章、節、…等の構成要素が文書の構造上の問題がないかをチェックするためのデータを生成して、テスト手段に渡し、テストが行なわれ、そのテスト結果に基づいて修正処理まで行うため、生成されたテストデータをチェックし、エラーがあれば自動的に修正を行うことが可能である。

【0022】第2の発明は、テストデータを生成する際に、入力された文書が構造化文書である場合には、構造上の各構成要素の内容を抽出して、どのような構成要素が文書内に位置しているかを構成要素毎に収集すると共に、当該構成要素毎に構造表示データと導入部を抽出して表示することにより、エラーを判定するための構成要素及びその構成要素毎の構造表示データ、導入部の内容をテストデータとして表示することが可能である。また、入力された文書が構造化文書以外であっても同様に段落を見つけることにより、テストデータを生成することが可能である。

【0023】第3の発明は、上記の第2の発明で表示された複数の構成要素の中から構造表示データまたは導入部を含めて表示させたい構成要素を利用者に指定させ、さらに、構造表示データのみを表示するのか、導入部も合わせて表示するのかを選択させることが可能である。従って、文書のテストを行う場合に、テストデータとして構造表示データのみと、さらに、当該構成表示データに対応する導入部も加えたテストデータとを選択して生成することが可能である。

【0024】第4の発明は、上記の第3の発明において、利用者が選択し、表示された構成要素毎の構造表示データや導入部分等を文書の構造情報と共に外部記憶装置等に記憶することが可能であるため、テストデータ生成処理が終了して、以降のテスト処理や修正処理を行う際に、そのまま移行しない場合、即ち、テストデータ生成処理のみを1つの処理として、完了させ、生成されたテストデータを記憶することにより、テストデータ生成処理をテスト処理や修正処理と切り離して行うことが可

能となる。テスト処理を行う場合には、テストデータ生成処理において生成されたテストデータを記憶手段より読み出して行うことが可能であると共に、随時用紙または、音声による読み上げも可能である。

【0025】第5の発明は、テスト処理に関する発明であり、入力されたあるいは、自動的に生成されるテストデータに構成要素の構成表示データの情報が含まれている場合には、構成要素間の階層関係をチェックする。また、テストデータに導入部の情報も含まれている場合には、当該構成要素の導入部の文章の最初の一文の主語が、同じ構成要素内で統一されているか、各導入部の文末の表現が文書内で統一されているかを調べることが可能となる。

【0026】第6の発明は、構成表示データの内容が所定のルールと合致しない場合には、ルールを参照する構成要素（階層）を上位または、下位に変化させることにより、構成表示データがどの構成要素に位置しているかの情報を取得することにより、単に所定のルールの構成要素と合致していないというだけでエラーとして出力しないため、テスト結果を表示し、修正する際に、修正する構成要素を特定することが可能である。

【0027】第7の発明は、構成表示データの項番がシーケンシャルに並んでいるかをチェックすることにより、順序通りに並んでいない場合には、エラーとして出力する。第8及び第9の発明は、テスト処理において、入力されたテストデータのすべての導入部の文末部分をチェックして、常体であるか敬体であるかを判断し、常体と敬体の個数を比較することにより、個数の多い方を導入部での「文末の形態」と見做す。そして、導入部間で「文末の形態」（常体か敬体か）を調べ、個数の少ない方を誤りとみなし、その誤りのを有する導入部を指摘する。また、全体の文書の導入部の文末形態から各々の導入部の文末の形態の正しさを判定し、エラーを指摘しておくことで、修正処理において正しい文末の形態に修正できる。

【0028】さらに、導入部内の文末に関しても、誤りの指摘を行う手段を有する。第10の発明は、導入部の

文書格遣表

要素ID	要素名	内容
1	章	要素ID: 2, 要素ID:3
2	段落	ツールの概要は以下になっています

上記のような文書構造表を生成する方法は、特開平 4-286054に記載されている方法を使用するものとする。この方法は、文書の文書構造を電子化された原稿から章、節、項、小項等の見出しや、図、表の見出しを抽出するという既知の方法を使用し、見出しに該当するも

最初の一文の主語を取り出して、単語に分解した結果、単語の数が2個以上ある場合であっても、当該主語毎に先頭の単語と最後の単語に細分化して、一致する単語のカウンタを取り、少ないものをエラーとする。

【0029】第11の発明は、テスト結果を出力する際に、表示手段に表示する方法と、外部記憶媒体等に記憶する方法がある。テスト処理に引き続いて修正処理を行う場合には、表示手段に表示すればよく、また、修正処理を後刻行う場合には、記憶媒体に格納されているテスト結果を使用すればよい。さらに、表示手段のみならず、随時用紙による出力や音声による読み上げも可能である。

【0030】第12の発明は、テスト結果に指摘すべきエラーの位置情報を含むように編集することにより、修正処理において、エラーの位置情報に基づいてエラーを検索し、適宜修正が可能となる。第13の発明は、修正処理に関するもので、テスト結果によりエラーとして指摘されている箇所を検索して修正する。テスト結果には、エラーとなっている原因が記載されているため、その原因に基づいてエラー箇所を修正することが可能である。

【0031】

【実施例】以下に、図面と共に本発明の実施例を詳細に説明する。図1は、本発明の一実施例の構造テスト装置の構成を示す。同図に示す構造テスト装置100は、テストデータ生成部10、文書格納部20、構造表示部30、抽出結果格納部40、構造テスト部50、テスト結果格納部60、修正部70、修正結果反映文書格納部80及び表示部格納部90より構成される。

【0032】テストデータ生成部10は、文書読取部11、文書構造検出部12、文書構造抽出部13より構成される。文書読取部11は、構造化文書や、段落で構成される一般の文書を読み込む。文書構造検出部12は、文書読取部11により入力された文書の文書構造を検出し、表1に示すような文書構造表を生成する。

【0033】

【表1】

のを全部抽出する。なお、この方法は、項番自体及び項番の情報を取り出すと共に、文書名、文書番号、文書種別（手引書類、解説書類等）、見出しの開始位置、見出しの全体の長さ、見出しの種別（部・章・節・項・小項・小見出し、図・表・付録等）項番の開始位置、項番の

長さ、項番の情報、見出しの内容の開始位置、見出しの内容の長さ、導入部データへのポインタを持つデータを作成する。なお、見出しの内容の長さが零の場合には、項番だけの見出しを意味する。

【0034】項番の情報には、記述されている項番の番号(例、1. 1. 1. 1であれば、1111とする。

(第1章の第1節の1項の1小項を表す))が入力される。また、図、表については、例えば、『第1図』である場合には1を、『表2』の場合には2を設定する。

『第1章』の場合には1を入力する。

【0035】なお、上記の例は、構造化文書の例であるが、一般の文書の場合には、項番がなく、第1章、第1節等の内容に基づいて文書構造表が生成される。文書構造抽出部13は、文書構造表の各要素より見出しに該当する部分と、導入部に当たる部分を抽出して構造表示部30に転送する。

【0036】文書格納部20は、作成されたマニュアル等の文書や、一般的な段落構成を有する文書を格納している外部記憶媒体である。構造表示部30は、テストデータ生成部10で生成された見出しデータ、導入部、及

び出力件数等を表示する。
【0037】抽出結果格納部40は、見出しデータ及び導入部データを抽出結果として格納する。また、表示部格納部90は、構造表示部30に表示されたデータ及び文書構造表の内容、見出しデータ及び導入部データを格納する。構造テスト部50は、テストデータ生成部10で生成されたテストデータまたは、表示部格納部90に格納されているデータを読み出す、または、構造表示部30に表示されているデータに基づいて構造的テストを行い、その結果をテスト結果格納部60に格納する。

【0038】修正部70は、構造テスト部50でのテスト結果に基づいて、必要であれば、修正を行い、修正した結果を修正結果反映文書格納部80に格納する。次に、上記の構成の一連動作を説明する。本実施例では、最初にテストデータ生成部10において、入力された文書に基づいて、階層毎の構成要素、構成要素の内容、構成要素の個数等のテストデータを生成し、次に構造テスト部50において、生成されたテストデータに基づいて文書の構造的な矛盾や、見出しの用語の誤り、導入部の統一性等のテストを行い、テストの結果、修正の必要がある場合には、修正部70において、自動的に修正を行う。

【0039】以下に、上記の各処理を説明する。

〔テストデータ生成処理〕図2は、本発明の一実施例のテストデータ生成部の動作の概要を説明するためのフローチャートである。

【0040】ステップ101) テストデータ生成部10の文書読取部11は、文書格納部20より、マニュアル等の文書を読み出して、文書構造検出部12に転送する。

ステップ102) 文書構造検出部12は、入力された文書に基づいて文書構造を検出する。文書構造としては、部、章、段落等の要素名と当該要素の内容等を検出し、文書構造表を生成する。文書構造表は、前述の表1に示すような内容により構成される。文書読取部11が読み込む文書はマニュアル等の構造化文書及び小説等の段落により構成される文書も含む。文書構造検出部12は、前述の表1に示すような文書構造表を文書構造抽出部13に転送する。

10 【0041】ステップ103) 文書構造抽出部13は、上記の文書構造表の各要素を参照して、文書構成に該当する部分(見出し部分:章、節、項等)と、導入部に当たる部分を抽出して構造表示部30に転送すると共に、抽出結果を抽出結果格納部40に格納する。

【0042】ステップ104) 上記で抽出された文書構成部分及び導入部を用いて、構成要素名とその個数を表示する。例えば、部の数が1、章の数が3、節の数が33、…等が抽出されているので、これらのデータを構造表示部30に表示する。また、同時に、各見出しや当該見出しに対応する導入部を表示する。これらのデータは、上記の表1の要素名欄より見出しを取り出し、内容欄より導入部の文を取り出して表示する。

【0043】ステップ105) 上記の構成要素及び個数の表示において、個数が0の構成要素の場合には、ステップ110に移行し、個数が1以上であれば、ステップ106に移行する。

ステップ106) 利用者が構造表示部30に表示されている構成要素からマウス等のポインティング装置を用いて取得したい構成要素を選択する。

30 【0044】ステップ107) ステップ104において、利用者が見出しのみを選択している場合には、ステップ108に移行し、見出し及び導入部の表示を希望している場合にはステップ109に移行する。

ステップ108) 構造表示部30は、構成要素と、当該構成要素に関する見出しを表示し、ステップ111に移行する。

【0045】ステップ109) 構造表示部30は、構成要素、見出し及び導入部を表示し、ステップ111に移行する。

40 ステップ110) 構造表示部30は、構成要素の個数が0の構成要素を利用者が選択しないように制御するか、または、選択した場合には、エラーメッセージ等を表示するようにしてもよい。

【0046】ステップ111) 上記で表示された内容を表示部格納部90に格納する。このとき、利用者により格納したい構成要素のみを構造表示部30に表示されている中から選択して格納してもよいし、全構成要素を格納してもよい。

次に、上記のテストデータ生成から構造表示までの具体的な動作を説明する。

【0047】まず、図3は、本発明の一実施例のテストデータ生成部の具体例を示す。同図の例において、図1と同一構成部分には、同一符号を付し、その説明を省略

『・・・第1章 ツールの概要は、・・・』 … a

であるとする。読み込まれた文書の内容 a は、文書構造検出部 12 に転送される。

【0048】(2) 文書構造検出部 12 は、上記の文書 a の内容の構造を検出し、図 3 b に示すような文書構造表を生成する。この文書構造表は、入力された文書には、要素 ID 2、要素 ID 3 及び要素 ID 5 を有する

(本来要素 ID は、シーケンスに並ぶものであるが(第1章、第2章、第3章、…)、少なくとも同図に示す例では、第1章と第4章が欠けている)。このような状態をテスト処理で指摘し、修正処理で修正される。さらに、同図 b の最初の段落の内容は、『ツールの概要は以下のようになっています』であることを示す。

【0049】(3) 次に、文書構造表 b を構造抽出部 13 に転送する。

(4) 構造抽出部 13 は、転送された文書構造表に基づいて、各構成要素を抽出する。ここで、抽出された構成要素はそのまま、抽出結果格納部 40 に格納される。抽出された構成要素を構造表示部 30 に転送する。

【0050】(5) 図 4 において、構造表示部 30 は、上記の図 3 の文書構造表 b の内容を表示する。構成要素として、表示される要素は、「部」、「章」、「節」、「項」、「小項」、「小見出し」、「図」、「表」、「付録」等である。これらの各要素毎に上記の

(1) で入力された入力文書内に存在する各項の数を文書構造表 b より取得して表示する。同図の例において、各構成要素の数は、部=1、章=3、節=33、項=50、小項=0、小見出し=0、図=15、表=0、付録=8、…となっている。

【0051】(6) 構造表示部 30 が上記の内容を表示すると、利用者はテストしたい構成要素を選択する。但し、個数が 0 の構成要素が選択された場合、即ち、図 4 の例において、“小見出し”が選択された場合には、構造表示部 30 は、『この構成要素には該当がありません。』等のエラーメッセージを表示させ、当該項目の選択を排他することも可能である。同図の例では、利用者は、『部』、『章』、『図』の 3 つの項目を選択している。さらに、見出しのみを表示するか、または、見出し及び導入部を表示するかも併せて選択させる。同図では、見出しのみを表示するパターンを選択している。

【0052】(7) 本例では、(6) の選択により見出しのみを選択しているため、同図 d に示すような内容が表示される。(6) において、選択されている構成要素は、『部』、『章』、『図』であるので、『部』の要素 ID から“第1部”とそのタイトル“導入手引き”、『章』の要素 ID から“第1章”とそのタイトル“導入に当たって”、『図』の要素 ID から“図 1. 1”とそ

する。

(1) まず、文書読取部 11 が文書格納部 20 より a に示すに入力文書を読み出す。この文書の内容は、

のタイトル“導入方法概略図”、『章』の要素 ID から“第2章”とそのタイトル“プロトコル”、…のように表示される。

【0053】また、上記の(6)において、利用者により図 4 に示す“見出し及び導入部”が選択された場合には、表示画面は、図 5 に示すように導入部も合わせて表示する。同図において、点線で囲まれている部分は、導入部の内容が表示されていることを示す。図 5 では、第2章のタイトルと導入部までを表示しているが、図 4 に示すように、部、章、図を選択した場合には、章は3番目までの章及び導入部が表示され、また、図は15番目の図の見出しまでが表示されることになる(図示せず)。

(8) 次に、構造表示部 30 に表示された内容及び文書構造、見出し、導入部等の情報を表示部格納部 90 に格納する。格納時に、利用者は、構造表示部 30 に表示されている構成要素を全て表示部格納部 90 に格納するようにしてもよいし、または、テストする必要があると思われる箇所を選択して格納してもよい。また、上記の(5)において、見出しのみが表示されている場合に、表示部格納部 90 に格納される要素は、導入部を含まない状態の情報が格納され、見出し及び導入部が表示されている場合には、見出し及び導入部の情報も併せて格納することができる。

【0054】つまり、見出しのみを格納する場合には、図 4 の d に示す内容及び文書構造情報が格納され、導入部も合わせて格納する場合には、図 5 に示される内容及び文書構造情報が格納される。なお、上記の(6)の処理において、図 4 及び図 5 の処理では、『部』や『章』を行頭を揃えて表示しているが、この例に限定されずに、階層構造がわかるように構成要素の種別毎に行頭をインデントさせて表示してもよい。図 6 は、階層構造を用いて表示した例を示している。『部』の行頭は左端から表示し、『章』の行頭は、左端から4文字分右にシフトさせて表示する。さらに、『図』及び『節』は『章』の行頭からさらに2文字分右にシフトさせて表示する。このようにすることにより、各構成要素毎の階層関係が明確になる。

【0055】[テスト処理] 次に、構造テスト部 50 は、構造表示部 30 に表示されているテストデータについてのテストを行うか、または、テストデータ生成部 10 により生成され、表示部格納部 90 に格納されているデータを読み出してテストするまたは、文書格納部 20 から直接入力された文書をテストすることも可能である。

【0056】構造テスト部 50 は、構成表示部 30 に表

示された情報、または、表示部格納部90から読み出した情報、入力された文書について、ある構成表示の下位に位置すべき見出しが上位に表示されていないか、反対に、上位に位置すべき見出しが下位に表示されていないか等を調べる。また、導入部に関しては、導入部の書き方が、同じ構成表示内で表現等が統一されているか否かを調べる。

【0057】図7は、本発明の一実施例の構造テスト部の動作を説明するための図である。構造テスト部50は、構造表示部30（または、表示部格納部90）から入力されたテストデータ、文書格納部20から直接入力された文書に対して、主語検査ファイル55、用語ファイル52、見出し定義ルールファイル53及び確認ファイル54を用いてテストデータに対するテストを行い、その結果をテスト結果表示部55に表示または、テスト結果格納部60に格納する。

【0058】構造テスト部50は、構造的なテストを以下の2点について行う。第1には、構造表示データ（見

『まず、この章では、2つの典型的な適用分野を紹介する。』…

と、という文があった場合、の文末“紹介する”は、修正の必要があるので、これらの結果をテスト結果として出力する。また、当然のことであるが、正しくない「文末の形態」をもっている導入部内の文末には、常体の文末をテスト結果として出力する。

【0061】用語ファイル52は、見出し定義ルールファイル53の見出しの語尾になる用語に該当しない語尾を持つ見出しを集めたもので、見出しのチェックの結果、エラーと見做せる見出しが格納される。不一致であった単語は、同じ文字種別（分類）の当該見出しの構成要素に対応する当該ファイル52内の部分（例えば、用語分類ファイル等を用意する）に出力する（例えば、章の「検査のための確認」という見出しに、定義ルールファイル53の同じ文書種別（分類）の内容に何も一致する用語がない場合には、用語ファイル52の“章”の構成要素を指すポインタ箇所に『「検査のための確認」』を追加する）。例えば、チェックした見出しの構成要素（例えば、章）、チェックした見出しの構成要素上での位置（章であれば、章中での位置…2番目の章であれば2とする）、該当見出しが属している文書種別（例えば、解説書類等）、見出し数のカウント（調べた見出しのうち用語ファイルに登録される見出しの数）が設定される。

【0062】見出し定義ルールファイル53は、各構成要素毎の見出しを定義するためのルールを保持する。図8は、本発明の一実施例の見出し定義ルールファイルの構成例を示す。これにより、構造テスト部50は、見出しの語尾が見出し定義ルールファイル53の「見出しの語尾となる用語」と一致しているか否かを調べる。同図（A）に示すように見出し定義ファイル53は、“手引き書類”、“解説書類”等の文書の分類情報と、各構成

出し）の上位になるべき見出しが下位に位置していないか、上位部分になるべき見出しが下位に位置していないか等の構成要素間の関係を検査する。

【0059】第2には、文書内の導入部の最初の一文の書き出しの部分と、全ての導入部間の文末の部分の書き方が統一されているかの検査を行う。例えば、第1章の導入部の最初の一文が、

『フロッピーディスクの構成は、……』…

であり、第2章の次に導入部の最初の一文が

『ここでは、……』…

である場合には、書き出しの単語として“フロッピーディスク”と“は”と、“ここでは”が取り出される。従って、及びの文の書き出し部分（主語）は、統一されていないことが分かる。

【0060】次に、構造テスト部50は、導入部の文末を検査する。導入部内及び導入部間で文末をチェックし、「文末の形態」として敬体が正しいと判断された場合、正しい「文末の形態」をもっている導入部内に、

要素、及び各構成要素に該当する用語より構成される。

また、同図（B）に示すように、木構造は、用語ファイル52に格納されている見出しから作成するもので、定義ルールファイル53の充実を図るために追加処理の参考にする場合や、オリジナルな定義ルールファイル53の作成をする時の参考とするものである。

【0063】確認ファイル54は、見出しの上位・下位が逆転されていないかを確認するためのファイルであり、見出し定義ルールファイル53に格納されている見出しの語尾になる用語と照合した結果、見出しの構成要素、見出しの定義ルールファイル53で構成要素の食い違いがないかを調べて確認するために使用される。

【0064】確認ファイル54に格納されているデータより、以下のような見出しのチェックが可能となる。確認ファイル54を参照して、見出しの構成要素cと見出し定義ルールファイル53のどの構成要素の見出しの語尾になる用語と一致したかdを比較して、cとdの構成要素により、当該構成要素が上位か下位かを調べ、例えば、下位概念のものが、ある構成要素より上位に位置している場合には、エラーと判定する。c=dの場合には、正しいと判定する。

【0065】上記ような比較をするために、以下のようなフラグを立てて、そのフラグにより上位であるか下位であるかを判定する。見出しの構成要素（例えば、章なら、1のフラグをたてる）、見出し定義ルールファイル53において、同じ文書種別（分類）のどの構成要素の単語と一致したのか（例えば、当該見出しが節の単語であれば2のフラグをたてる）、調べた見出しの構成要素上の位置（章なら章中での位置、2番目の章なら2とする）、例えば、解説書類と、該当見出しの属している文書種別（分類）、見出し数のカウント（=1）等を設定

する。

【0066】主語検査ファイル55は、導入部のテストに用いられるファイルであり、各導入部の先頭の一文の主語を取り出し、その中の主語を抽出する。この抽出された主語を単語に分解し、単語数が2以上であれば、先頭と最後の単語を取り出して、検査する。単語数が1である場合には、当該単語を取り出す。(例えば、ここでは、“この事では”が主語の場合には、『この』と『では』を取り出す。このように取り出された単語が主語検査ファイル55に格納される。また、検査時に当該ファイル55が読み出される。

【0067】常体・敬体検査ファイル56は、導入部の文末を抽出したあとに当該文末が常体か敬体かを判定した判定結果が格納されるファイルである。次に、構造テスト部50の処理として、見出しのテストの動作を説明する。図9は、本発明の一実施例の構造テスト部の見出しのテストの概要の動作を示すフローチャートである。

【0068】ステップ201) 次のデータが存在する場合には、ステップ202に移行し、最終のデータの処理が終了したら、テスト処理を終了する。

ステップ202) 文書格納部20より直接文書が入力される、または、構造表示部30から(または、表示部格納部90からデータを読み出して)データが入力されると、構造テスト部50は、最も上位の見出しから順にテストを行う。つまり、最も上位の構成要素が『部』であれば、『部』のテストが全て終了したら、次の『章』のテストに移行する。

【0069】まず、入力された見出しを順に取り出す。ステップ203) 取り出した見出しを項番と内容に分ける。

ステップ204) 項番だけの場合には、ステップ213に移行する。

ステップ205) 見出しの用語について、見出し定義ルールファイル53のデータを読み込み、照合を行う。また、項番がシーケンシャルに並んでいるかをチェックする。項番のチェック方法については、後述する。

【0070】ステップ206) 見出し定義ルールファイル53を読み出す際に、見出し定義ルールファイル53のどの分類を使用するかを決定する。なお、文書分類が「その他」が検出された場合には、ステップ210に移行する。次に、見出し定義ルールファイル53と見出しの用語が一致するかを検査する。調べる見出しの構成要素の内容の語尾が見出し定義ルールファイル53の見出しの語尾になる用語に一致しているかを検査する。

【0071】ステップ207) 見出し内の用語の語尾が見出し定義ルールファイル53の見出しの語尾になる用語と一致している場合には、確認ファイル54に情報を出力する。確認ファイル54には、調べた見出しの構成要素名、一致した見出し定義ルールファイル52の構

成要素、文書種別、構成表示上の位置情報、見出し数のカウント等を記録し、ステップ201に移行する。確認ファイル54には以下のような内容が出力される。

調べた見出しの構成要素(例えば、“章”のとき=1)：

見出し定義ルールファイル53のどの構成要素の単語と一致したか(例えば、“節”のとき=2)：

調べた見出しの構成要素上の位置(章なら章中での位置、2番目の章であれば=2)：

10 該当見出しの属している文書種別(解説書類=1)：
上記の～によりデータを生成すると、“1221”という1つのデータになる。

【0072】ステップ208) 見出し定義ルールファイル53の見出しの語尾になる用語と見出し内の語尾の用語が一致しない場合には、構成要素の階層を変えて調査する。詳細は、後述する。

20 ステップ209) 全構成要素についてチェックが終了すればステップ210に移行し、まだ、他の構成要素のチェックが終了していなければ、ステップ206に移行する。

【0073】ステップ210) 用語ファイルに調べた見出しの構成要素、調べた見出しの構成要素上の位置、文書種別、見出しの用語、用語ファイルに出力した見出しの数を出力し、ステップ201に移行する。

ステップ211) ステップ204において、見出しが項番のみの場合には、当該項番の情報を見出しの構成要素、調べた見出しの構成要素上の位置エラーファイル(図示せず)に出力する等のエラー処理を行い、ステップ201に移行する。

30 【0074】上記のステップ205における項番のチェックについて説明する。

最初に項番の情報を取り出す。

前述の表1に示すような文書構造表を参照しながら、“章”内であれば章内にある項番(節、項、小項等、または、図表の項番等)の検査を行う。節なら節が一つずつインクリメントされているか、項なら項がインクリメントされているか等を検査する。このとき、一番上位の見出しが1から始まっているか、いない場合には、いくつかから始まるかは利用者に指定させることも可能である。このため、以下のような2つの検査が可能となる。

(1) 最上位の構成要素の見出しを除いて構成要素が1から始まっているか(最上位が“章”であれば、“節”、“項”等がその章内で1から始まっているか)。

(2) 項番が順番通りになっているか(飛びや逆転現象等になっていないか)。

【0075】具体的には以下のようにチェックする。例えば、“1. 1. 1. 1”なら“1111”の数値に置き換える。また、“付録a. 1”という場合には、

“a”を“1”等の数値に置き換え、“11”とする等の方法が可能である。このような変換を行うことにより、数値化して前の項番の数値に1をインクリメントすると、次に項番の数値になるか否かによりチェックすることが可能である。但し、これは、同じ構成要素内で行われるものとする。

【0076】例えば、
“1. 1. 1. 1. 99”

$$(11199-99) \times 10 + (99+1) = 111100 \quad (1)$$

上記のようにして等しければ、正しいと判定し、等しくなければ誤りと判定する。上記の式(1)の“99”は、調べる項番により変化するが、この場合には、“1. 1. 1. 99”の場合二桁であるためであり、三桁であれば、“999”となり、一桁であれば“9”となる。

【0078】図10は、本発明の一実施例の用語ファイルに構成要素毎の用語を出力する動作のフローチャートである。

ステップ2091) ステップ206において、現在の構成要素が「節」である場合に、「節」より下の構成要素「項」で見出しでいざルールファイル53の見出しの語尾になる用語と一致する用語がないか調べる。一致する場合には、ステップ207の処理に移行する。不一致の場合には、ステップ2092に移行する。

【0079】ステップ2092) それでも全ての下位方向の構成要素で一致しない場合には、「項」よりさらに、下位の構成要素「小項」で一致する単語が見出し定義ルールファイル53にないかを調べる。一致する場合には、ステップ207の処理に移行し、不一致の場合には、ステップ2093に移行する。

【0080】ステップ2093) 上記のステップ2092において、最下位の構成要素「小項」でも見出し定義ルールファイル53と一致しない場合には、次に「節」より上位の構成要素である「章」で一致する用語が見出し定義ルールファイル53にないかを調べる。一致する場合には、ステップ207に移行し、不一致の場合には、ステップ2094に移行する。

【0081】ステップ2094) 上記のステップ2093において、一つ上位の構成要素でも不一致の場合には、さらに上位の構成要素「部」で一致する用語が見出し定義ルールファイル53に存在しないかを調べる。一致する場合には、ステップ207に移行し、不一致の場合には、ステップ2095に移行する。

【0082】ステップ2095) 上記のいずれの構成要素であっても見出し定義ルールファイル53のデータと一致するデータがない場合には、用語ファイル52に調べた見出しの構成要素、調べた見出しの構成要素上の位置を文書種別、当該用語及び見出しカウント数をインクリメントして書き込む。

【0083】次に、具体例を説明する。

の次の項番が

“1. 1. 1. 1. 100”

で正しいことを検査するために、

“1. 1. 1. 99”

を数値化すると、11199であり、

“1. 1. 1. 100”

を数値化すると、111100となる。

【0077】

1. 『第1部』

(1-1) まず、最初に見出しの構成要素『部』が入力される。

(1-2) 次に、上記の『第1部』については、項番だけの見出しであるので、エラー情報(見出しの構成要素、調べた見出しの構成要素上の位置)を出力する。

【0084】2. 『第1章 画像処理の全体像』

(2-1) 次に構成要素『章』及び見出し“画像処理の全体像”が入力される。(2-4) 構造テスト部50は、見出し定義ルールファイル53には、構成要素“章”に関する見出しの語尾になる用語として、上記の見出しの語尾「全体像」が一致するため、確認ファイル54に見出しの構成要素、見出し定義ルールファイル53と一致した構成要素、見出しの構成要素上の位置、該当見出しの属している文書種別、見出し数のカウント値をインクリメントして書き込む。

【0085】3. 『図2. 1』

(3-1) 上記で『節』のチェックが終了したので、第1章に含まれる次の構成要素『図』、及び項番“2. 1”が入力される。

30 (3-2) 構造テスト部50は、上記の見出しが項番だけの見出しと判断し、エラー情報を出力する。

【0086】上記の処理により、用語ファイル52と確認ファイル54に格納されているデータを以下のような方法でチェックする。まず、確認ファイル54を参照して、誤りと見做す見出しについてエラーメッセージをテスト結果表示部57に出力する。なお、確認ファイル54になにもデータが格納されていない状態は、見出し定義ルールファイル53に従っていないことになるため、エラーとするが、全ての見出しがエラーであるという指摘では、意味がないため、用語ファイル52の情報に基づいて、図8(B)に示す木構造を作成し、エラーを含む見出し誤り候補として出力することも可能である。また、木構造にシソーラスファイルを適応して誤りをチェックすることも可能である。

【0087】見出し定義ルールファイル53の見出しの語尾になる用語と一致しない場合には一応エラー対象と考え、用語ファイル52に格納された見出しをカウントし、用語ファイル52に出力しているため、最終的に用語ファイルの見出し数=0であれば、調べた見出しは、

50 見出し定義ルールファイル53の内容に従っていると考

えられる。また、用語ファイルの見出し数>0であれば、用語ファイル52の見出しを誤り候補とし、エラーメッセージとしてテスト結果表示部57に出力する。また、項番だけの見出しのエラーと項番の誤りを検査するため、これらのエラーを出力する場合には、構成要素と文書内の構成要素中の位置情報も併せて出力する。

【0088】このとき、確認ファイル54とデータ数が1以上である場合には、確認ファイル内の見出しの構成要素Cと見出し定義ルールファイル52のどの構成要素が見出しの語尾になる用語と一致したかdを調べ、c=dでない場合、誤りと見出し、エラー情報を出力する。また、確認ファイルの構成要素上に位置と項番の誤り及び項番だけの見出しのエラー情報に出した構成要素上の位置から抜けている構成要素上の位置の番号が誤りと見做せ、用語ファイル52に、その番号（抜けている構成要素上の位置番号）は入っているの、用語ファイル52を参照して見出しの内容に関するエラー情報を出力する。

【0089】次に、構造テスト部50における導入部のテストの動作を説明する。図11は、本発明の一実施例の構造テスト部の導入部のテストの概要動作を示すフローチャートである。

ステップ310) まず、構造テスト部50は、導入部のテストを行う際に、導入部の先頭の1文の主語を抽出する。

【0090】ステップ320) 導入部の主語が統一されているかを判定する。

ステップ330) 導入部の主語の処理が終了すると、次に全導入部文末を抽出する。

ステップ340) 導入部毎に抽出された文末が常体であるか敬体であるかを判定し、各々の数をカウントして、数の少ない方がエラーの可能性があると、抽出する。

【0091】図12は、本発明の一実施例の構造テスト部の導入部のテストの詳細動作を示すフローチャートである。

ステップ311) 構造テスト部50は、導入部の先頭の一文の主語を抽出する。抽出する方法は、TDS/MAPLE等の既存の方法を使用するものとする。

【0092】ステップ312) 抽出した主語を分解する。

ステップ313) ここで、主語に該当する単語数が2つ以上であれば、先頭と最後の単語を取り出す。もし単語数が1であれば、当該単語を取り出す。以下に単語分解の例を示す。

【0093】まず、区切り記号を含まず、かつ、同じ文字種が続く一連の文字列を抽出する。但し、数値は、漢字であっても区切り記号を見做す。区切り記号は“[]”や“{}”及び“”、“,”、“.”等や空白を含む。また、数値が出現した場合には、区切り記

号とみなし、数値は含めない。

【0094】また、本例では、どの構成要素に対応する導入部の主語を取り出したのかの情報を記憶しておくものとする。

例1) 『この章では、システムの概要について述べます。』

主語の抽出：この章では

単語分解：「この」「章」「では」（単語数3）

主語検査ファイルに書き出す情報：導入部の構成要素／章

単語数／3

調べる単語／先頭「この」

最後「では」

例2) 『第3章においては、システムの体系について記述しています。』

主語の抽出：第3章においては

単語分解：「第」「章」「においては」（単語数3）

主語検査ファイルに書き出す情報：導入部の構成要素／章

単語数・3

調べる単語／先頭「第」

最後「においては」

例3) 『6. 10. 8では、マウスの使い方について説明します。』

主語の抽出：6. 10. 8では

単語分解：「では」（単語数1）

主語検査ファイルに書き出す情報：どの導入部か／項
単語数／1

調べる単語／先頭「では」

最後 なし

例4) 『この9. 12節では、システムのメリットについて説明します。』

主語の抽出：この9. 12節では、

単語分解：「この」「節」「では」（単語数3）

主語検査ファイルに書き出す情報：どの導入部か／節
単語数／3

調べる単語／先頭「この」

最後「では」

例5) 『ここでは、システムの概要について述べます。』

主語の抽出：ここでは

単語分解：「ここでは」（単語数1）

主語検査ファイルに書き出す情報：どの導入部か／部
単語数／1

調べる単語／先頭「ここでは」

最後 なし

ステップ314) ステップ313で取り出された各単語は主語検査ファイル55に出力する。

【0095】上記のステップ311から314の処理をすべての導入部に対して行う。図13は、本発明の一実

施例の構造テスト部の導入部のテスト動作を示すフローチャートである。

ステップ 3 2 1) 次に、構造テスト部 5 0 は、主語検査ファイル 5 5 を読み出す。

【0096】ステップ 3 2 2) 同じ構成要素の導入部を取り出す。

ステップ 3 2 3) 読み出した主語検査ファイル 4 5 5 の単語数 = 1 の主語と単語数 = 2 以上の主語の数をカウントし、同じ構成要素の導入部に関する調査が終了したら、ステップ 3 2 2 に移行する。

【0097】ステップ 3 2 4) 単語数 = 1 と単語数 = 2 以上の数が両方共に 1 以上であるとき、ステップ 3 2 5 に移行し、それ以外の場合には、ステップ 3 3 0 に移行する。

ステップ 3 2 5) 単語数 = 1 の主語と単語 2 以上の主語では明らかに相違があるので、まず数だけの比較をする。

【0098】ステップ 3 2 6) 単語数 = 1 の主語の数 > 単語数 = 2 以上の主語の数であれば、ステップ 3 3 2 に移行する。

ステップ 3 2 7) 単語数 = 1 の主語の数 < 単語数 = 2 以上の主語の数であれば、ステップ 3 2 8 に移行し、それ以外であれば、ステップ 3 3 5 に移行する。

【0099】ステップ 3 2 8) 単語数 = 1 の主語をもつ導入部を誤りと見做す。

ステップ 3 2 9) エラーを指摘する。

ステップ 3 3 0) 単語数 = 1 の主語の数が 0 である場合には、ステップ 3 3 1 に移行し、それ以外である場合単語数 = 2 以上の主語の数が 0 である場合には、ステップ 3 3 4 に移行する。

【0100】ステップ 3 3 1) ステップ 3 2 8 において、導入部を誤りと見做すと共に、単語数 = 2 以上の主語の検索を行う。詳細は後述する。

ステップ 3 3 2) ステップ 3 2 6 において、単語数 = 1 の主語の数 > 単語数 = 2 以上の主語の数であれば、単語数 = 2 以上の主語をもつ導入部を誤りとみなす。

【0101】ステップ 3 3 3) エラーの指摘を行う。

ステップ 3 3 4) さらに、単語数 = 1 の主語の検査を行う。また、ステップ 3 3 0 において、単語数 = 1 の主語の数 ≠ 0 である場合にもこの検査を行う。

ステップ 3 3 5、ステップ 3 3 6) ステップ 3 2 7 において、単語数 = 1 の主語の数 = 単語数 = 2 以上の主語の数となっている場合には、導入部の書き方がばらばらであるとの指摘を出し、図 1 7 の書き出しの相違の部分に、総ての導入部の主語の部分をページ、行番号も付与して出力して、注意を促す。なお、この部分については自動修正は不可能となる。

【0102】次に、上記のステップ 3 3 4 の単語数 = 1 の主語の検査について説明する。図 1 4 は、本発明の一実施例の単語数 = 1 の主語の検査方法を示すフローチャ

ートである。

ステップ 3 3 4 1) 主語検査ファイルに記憶されている単語数 = 1 で「調べる単語」の“先頭”の部分を持つだけ取り出す。

【0103】ステップ 3 3 4 2) 同じ語があるかを検査し、同じ語があったらカウントする。

ステップ 3 3 4 3) カウント数の多い順に並べる。

ステップ 3 3 4 4) カウント数の最も多いものを正しい書き出し（導入部）とみなし、他のカウント数の少ない語を持つ主語を誤り可能性ありと見做す。その結果、図 1 7 に示すように出力する。

【0104】図 1 5 は、本発明の一実施例の単語数 = 2 以上の主語の検査のフローチャートである。

ステップ 3 3 1 1) 主語検査ファイルに格納されている「単語数 = 2 以上」の「調べる単語」の“先頭”、“最後”の部分を持つだけ取り出す。

【0105】ステップ 3 3 1 2) 同じ形式の主語であるかを検査し、同じ形式の主語があったら、その形式をカウントする。但し、単語数 = 2 以上で構成される主語の場合、調べる単語として、「先頭」と「最後」の 2 つがあるので、以下の場合が考えられる。

【0106】「先頭」の単語にも一致し、「最後」の単語にも一致する。

例) 上記の例 1 と例 4 については同じ形式の主語とみなし、この形式でカウントする。

「先頭」の単語が一致しないで、「最後」の単語には、一致する。この場合、「先頭」の単語の文字種がひらがなの場合には、同じ形式であると思えない。他の文字種（漢字または、カタカナ、アルファベット等）の場合には、固有名詞の違いだけでも考えられるので、同じ形式の主語の可能性のあるものとし、可能性有りという認識のできる形でこの形式をカウントする。

【0107】例)

a. 利用の気体の状態方程式とは

単語数 4, 「理想気体」「の」「状態方程式」「とは」

先頭の単語「理想気体」

最後の単語「とは」

b. ポテンシャルとは

単語数 2 「ポテンシャル」「とは」

先頭の単語「ポテンシャル」

最後の単語「とは」

c. これらの電子雲とは

単語数 3 「これらの」「電子雲」「とは」

先頭の単語「これらの」

最後の単語「とは」

a と b は同じ形式の主語の可能性ありと見做せる。また、a と c は同じ形式の主語とは見做さない。

【0108】「先頭」の単語が一致して「最後」の単語が一致しない場合、この場合は、違う形式の主語で

あると見做す。

「先頭」と「最後」の単語両方が一致しない場合には、違う形式の単語であると見做す。

【0109】ステップ3313) カウントの多い順に並べる。

ステップ3314) カウント数の一番多いものを正しい書き出しと見做す。他のカウント数の少ない語を持つ主語を誤り可能性ありと見做し、図17のような指摘を出力する。

【0110】図16は、本発明の一実施例の導入部の文末のテストの動作を示すフローチャートである。

ステップ401) 導入部の文末を抽出する。この文末抽出方法は、特開昭62-78674及び特開平01-106161等に記載されている方法を用いるものとする。

【0111】ステップ402) 次に、抽出された文末が常体か敬体を判定する。この判定方法は、上記の特開昭62-78674及び特開平01-106161等に記載されている方法を用いるものとする。

ステップ403) 判定した結果を当該導入部の位置情報と共に、常体・敬体検査ファイル56に格納する。

【0112】ステップ404) 次に、構造テスト部50は、常体・敬体検査ファイル56を読み出す。

ステップ405) 読みだされた常体・敬体検査ファイル56の内容に敬体と常体が混在しているかを判定し、混在していない場合には、ステップ407に移行し、混在している場合には、ステップ406に移行する。

【0113】ステップ406) 文末の形態がそのまま決まるため、ステップ408に移行し、導入部の値を追加する。

ステップ407) 常体と敬体が混在している場合には、多い方をこの導入部での文末の形態とみなし、常体を形態検査ファイルに値を追加する。

【0114】ステップ408) 導入部全ての検査が終了したら、「文末の形態」の少ない方を誤りと見做す。上記のようにして、構造テスト部50でテストされた結果がテスト結果表示部57により、図17に示すよう見出し部に関しては用語ファイル52、確認ファイル54また項番のエラー情報を読み込み、導入部に関しては、各結果の数が少ない要素をエラーとして表示する。

【0115】さらには、テスト結果表示部57に表示された内容を外部記憶装置のテスト結果格納部60に格納することも可能である。テスト結果表示部57は、図17(A)に示すように、詳細に表示したい誤り箇所を利用者に指定させる。利用者は見出しや、導入部の何れかの内容をマウス等でクリックして指定する。図17

(A)の表示画面より、利用者が「見出し：ふさわしくない見出し」を選択すると、見出し定義ルールファイル53を当該エラーの可能性があるとされる要素にて検索し、図17(B)のaに示すように、当該見出しが存在

している位置とその理由を表示する。

【0116】また、図17(A)の表示画面より、利用者が「導入部：書き出しの相違」を選択すると、図17(B)のbに示すように当該第1の書き出し位置と第2の書き出し位置と、それぞれの導入部の内容を表示する。また、図17(A)の表示画面より、利用者が「導入部：文末の乱れ」を選択すると、図17(B)のcに示すように、文末が常体・敬体の統一がとれていない文の位置と内容を指摘する。

10 【0117】なお、図17において、「」内に検査した結果を挿入するが、この内容は、検査の結果により可変である。また、指摘箇所をマウス等のポインティング装置でポイントすることにより、当該本文の該当箇所に位置付けすることも可能である。このため、必要に応じて当該箇所にテスト結果が設定され、表示された段階で、利用者自身で修正することは勿論可能である。

【0118】[修正処理]次に、修正部70について説明する。修正部70は上記の構造テスト部50により出力されたテスト結果に基づいて自動的に修正を行う。

20 【0119】図18は、本発明の一実施例の修正部を説明するための図である。修正部70は、テスト結果表示部55に表示される図18(A)より修正したい箇所をマウス等でクリックする等して指定する。これにより、指定された箇所に対応するエラーの可能性を示す内容が図18(B)に示すように表示される。つまり、見出しに関しては、概念位置の乱れが3件あるので、図18(B)のaには、3件分のエラーの可能性のある箇所が表示される。また、図18の(B)のbには、導入部の書き出しの相違が8件あるため、8件分のエラーの可能性のある箇所がテスト結果表示部55に表示される。

30 【0120】修正部70は、図18(B)において、利用者が修正したい部分をクリックすると図19に示すように修正候補が表示される。この中より修正に適するものを選択することにより修正される。例えば、図18(B)aにより指摘されている492ページ、18行に記載されている文の位置をクリックすると、図19に示すような修正候補が表示される。利用者は図19に表示される3つの修正候補の中から置き換えたい候補の位置にカーソルを移動させ、マウスでクリックすることにより、自動的に修正される。

40 【0121】また、図18(B)cにより指摘されている115ページ50行に記載されている文の位置をクリックすると、図20に示すように自動的に常体から敬体に修正されて出力される。また、文字の乱れに関しては、複数の指定が可能であり、また、一括して自動修正することもできる。

50 【0122】このように、導入部の書き出しの相違、文末の乱れ等に関しては、修正指示として、所望の修正位置にカーソルを移動させて、クリックして指示することにより、図20に示すように、当該位置についてルール

に沿った修正を行う。また、構造テスト部50において、テスト結果を書き込む確認ファイル54や用語ファイル52をファイルとして持たずにメモリ内に保持してもよい。

【0123】なお、上記の構造表示部30やテスト結果表示部55は、図面及び説明上分けて記載しているが、物理的には、同一のディスプレイを用いるものとする。なお、本発明は、上記の実施例に限定されることなく、特許請求の範囲内で種々、変更、応用が可能である。

【0124】

【発明の効果】第1の発明によれば、文書により任意の要素を簡単に取り出し、表示できるようになる。また、文書の構造的なテスト及び修正を自動的に行うことが可能となり、修正作業は、表示された内容の単に確認作業を行えばよく、利用者の操作負担が軽減される。

【0125】第2の発明によれば、作成済または作成途中の文書の文書構成上の階層や、見出し及び導入部分等を抽出しテストデータを生成することにより、以降のテスト処理におけるデータ取得が容易になる。第3の発明によれば、作成された構成要件の件数と表示させるデータが見出しのみか、導入部も含めて表示するかを利用者に選択させる。これにより、利用者は、詳細に見たい構成要件と見出し及び導入部の確認が可能である。また、表示された内容がそのままテスト処理に移行されるので、利用者は、引き続いてテストを実行することができる。

【0126】第4の発明によれば、第3の発明により表示されているテストデータを外部記憶装置等の記憶装置に保持することが可能であるため、第3の発明によりテストデータが生成され、表示された段階で生成処理が終了しても、記憶しておくことにより、後日、テスト処理を行う場合に、記憶装置に保持されているテストデータを用いて処理することも可能である。さらに、用紙等の他の媒体にも出力することが可能となる。

【0127】第5の発明によれば、各構成要件と当該構成要件の件数が表示されている画面よりテストを行いたい構成要件を利用者に選択させることが可能であり、選択された構成要件の見出しの位置関係をチェックし、さらに、導入部の文が統一された表現になっているか等のテストを行うことが可能である。

【0128】第6の発明によれば、見出し部分の構成要素が所定のルールと合致しない場合でも、構成要素を上下させてルールと照合することにより、修正時にどの構成要素に修正すればよいかを把握することが可能である。第7の発明は、項番に関してもシーケンシャルに並んでいるか否かを判定するため、重複、飛び等があってもそれらのエラーを把握でき、修正時にも自動的に修正することが可能である。

【0129】第8の発明は、導入部分のチェックを文書全体の導入部の間で行うため、修正時には、一つの文書

を通して統一された導入部に修正することも可能である。第9の発明によれば、主導入部の接尾語を敬体または常体かを判定し、導入部全体の文末形態を判断し、個数の少ない方をエラーの可能性のあるものとして指摘するため、統一すべき用語を修正時に選択することが可能である。

【0130】第10の発明によれば、統一すべき主語を正確に選択することが可能である。第11の発明によれば、テスト結果を表示または、外部記憶手段に出力することにより、そのまま、修正処理に移行することが可能である。また、外部記憶手段に格納したデータを音声情報に変換したり、用紙等に出力することも可能である。

【0131】第12の発明によれば、テスト結果にエラーとされた位置を明示しておくことにより、修正時には当該エラー位置を特定することが可能である。第13の発明によれば、テスト結果において、エラーと判断された箇所を自動的に修正することが可能である。

【図面の簡単な説明】

【図1】本発明の一実施例の構造テスト装置の構成図である。

【図2】本発明の一実施例の動作の概要を示すフローチャートである。

【図3】本発明の一実施例のテストデータ生成部の具体例を示す図である。

【図4】本発明の一実施例の構造表示部の具体例を示す図（その1）である。

【図5】本発明の一実施例の構造表示部の具体例を示す図（その2）である。

【図6】本発明の一実施例の構造表示部の表示例を示す図である。

【図7】本発明の一実施例の構造テスト部の動作を説明するための図である。

【図8】本発明の一実施例の見出し定義ルールファイルの構成例を示す図である。

【図9】本発明の一実施例の構造テスト部の見出しのテストの概要の動作を示すフローチャートである。

【図10】本発明の一実施例の用語ファイルに構成要素毎の単語を出力する動作のフローチャートである。

【図11】本発明の一実施例の構造テスト部の導入部の概要動作を示すフローチャートである。

【図12】本発明の一実施例の構造テスト部の導入部の詳細動作を示すフローチャートである。

【図13】本発明の一実施例の構造テスト部の導入部のテストの動作を示すフローチャートである。

【図14】本発明の一実施例の単語数＝1の主語の検査方法を示すフローチャートである。

【図15】本発明の一実施例の単語数＝2以上の主語の検査方法を示すフローチャートである。

【図16】本発明の一実施例の導入部の文末のテストの

動作を示すフローチャートである。

【図 1 7】本発明の一実施例のテスト結果を示す図である。

【図 1 8】本発明の一実施例の修正部の動作を説明するための図である。

【図 1 9】本発明の一実施例の修正部における修正候補表示の例を示す図である。

【図 2 0】本発明の一実施例の文末の乱れの修正例を示す図である。

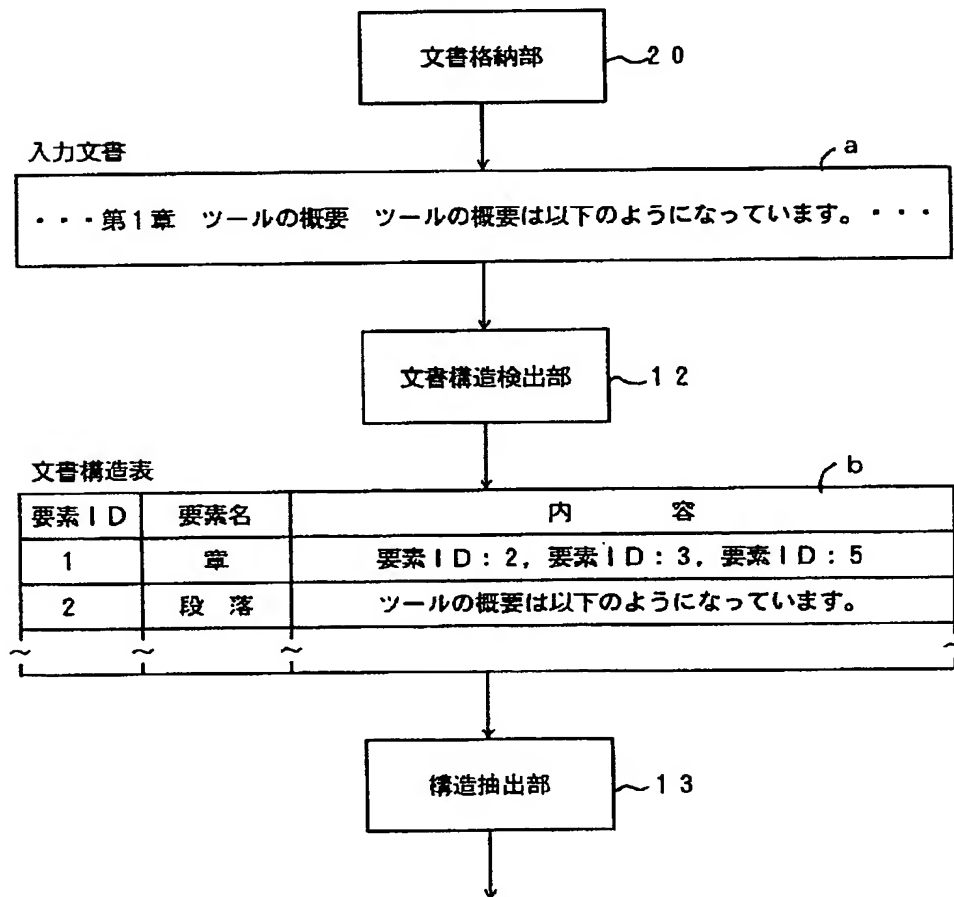
【符号の説明】

- 1 0 テストデータ生成部
1 1 文書読取部
1 2 文書構造検出部
1 3 文書構造抽出部
2 0 文書格納部

- 3 0 構造表示部
4 0 抽出結果格納部
5 0 構造テスト部
5 2 用語ファイル
5 3 見出し定義ルールファイル
5 4 確認ファイル
5 5 主語検査ファイル
5 6 常体・形態検査ファイル
5 7 テスト結果表示部
10 6 0 テスト結果格納部
7 0 修正部
8 0 修正結果反映文書格納部
9 0 表示部格納部
1 0 0 構造テスト装置

【図 3】

本発明の一実施例のテストデータ生成部の具体例を示す図



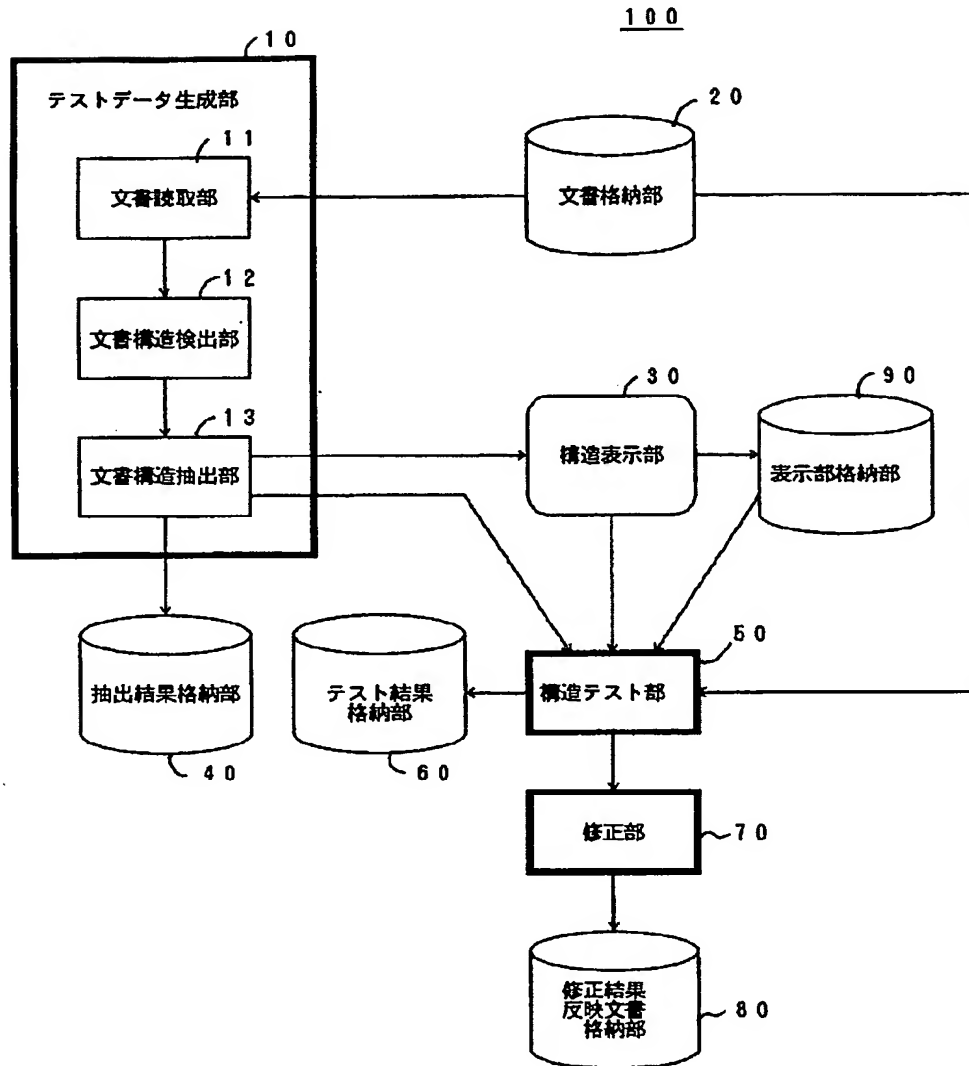
【図 2 0】

本発明の一実施例の文末の乱れの修正の例を示す図

例) 第2章では
↓
ここでは、
↓
.....であった。
↓
.....です。

【図 1】

本発明の一実施例の構造テスト装置の構成図



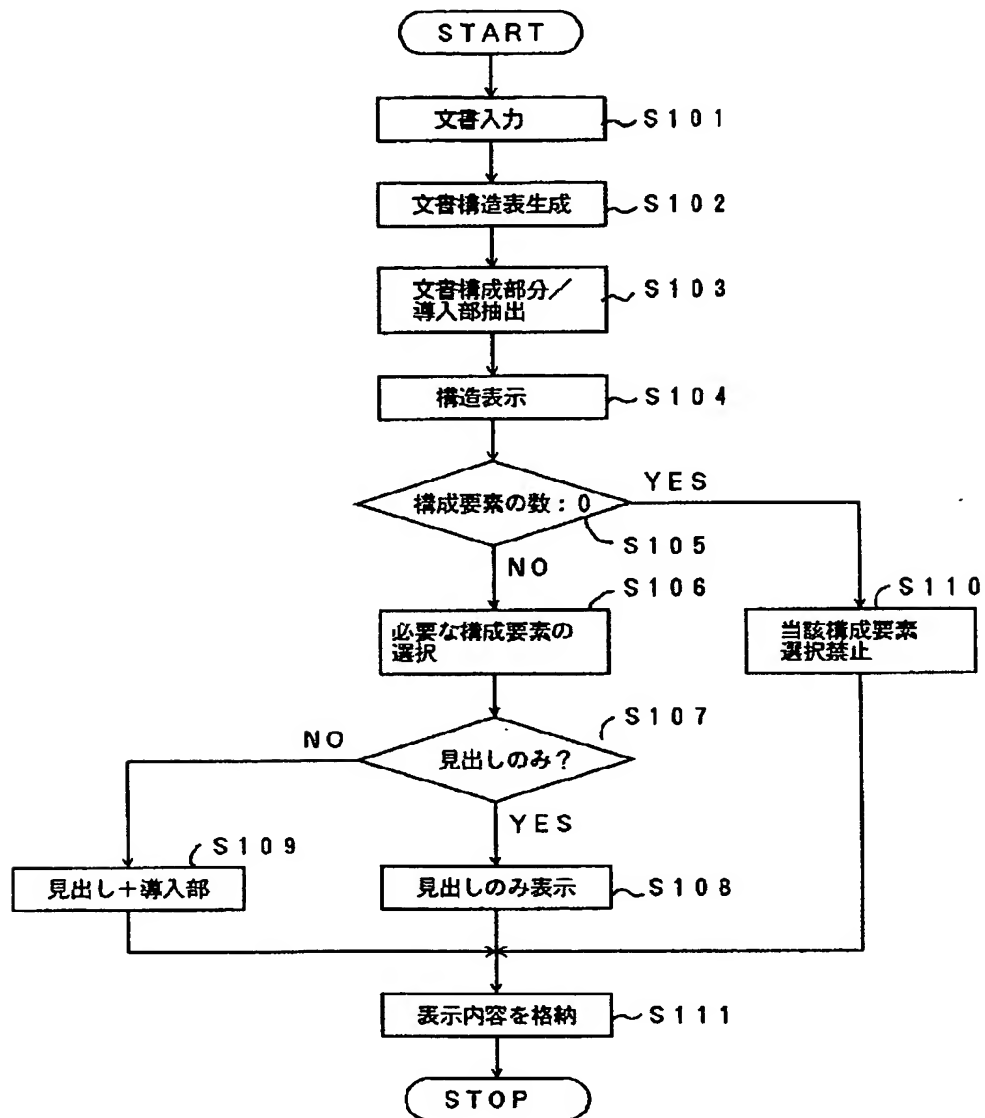
【図 19】

本発明の一実施例の修正部における修正候補表示の例を示す図

第3章 TDS/MAPLEの全体像	↑
第3章 TDS/MAPLEの特長	
第3章 TDS/MAPLEの機能概要	
⋮	
↓	

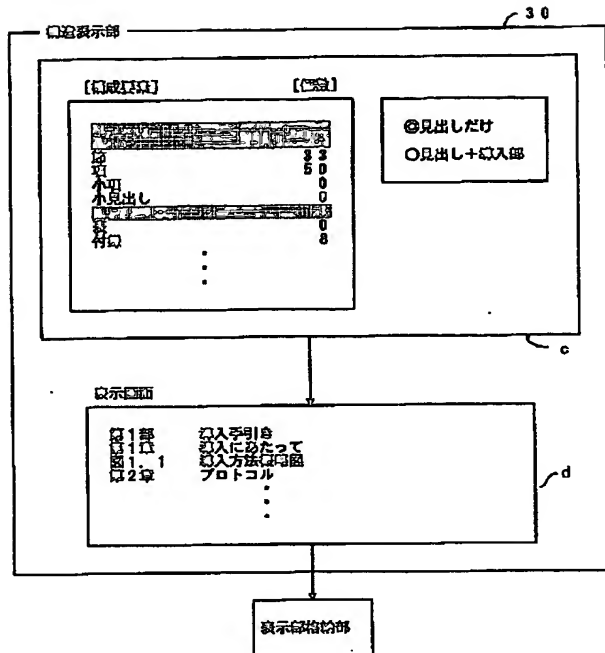
【図2】

本発明の一実施例のテストデータ生成部の動作の概要を示すフローチャート。



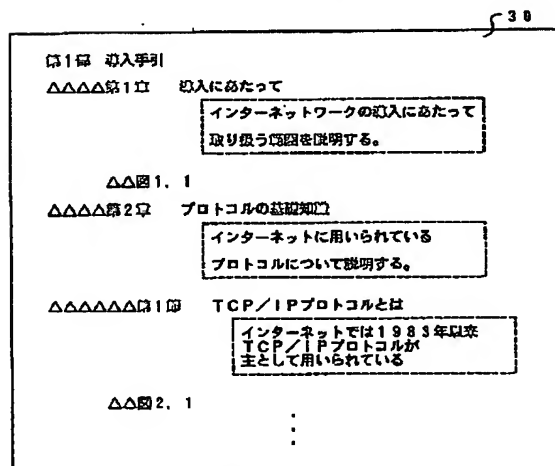
【図4】

本発明の一実施例の画面表示部の具体例を示す図（その1）



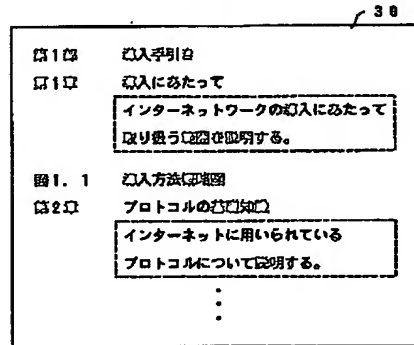
【図6】

本発明の一実施例の検索表示部の表示例を示す図



【図5】

本発明の一実施例の画面表示部の具体例を示す図（その2）



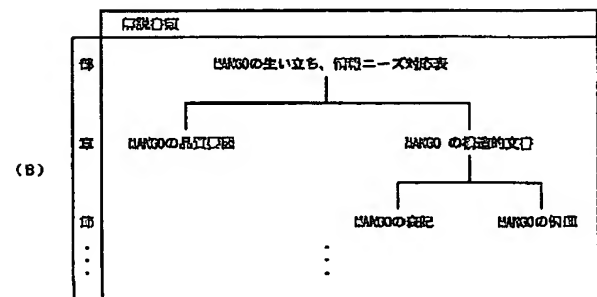
【図8】

本発明の一実施例の見出し定口ルールファイルの和成例を示す図

見出し定口ルールファイルの一例

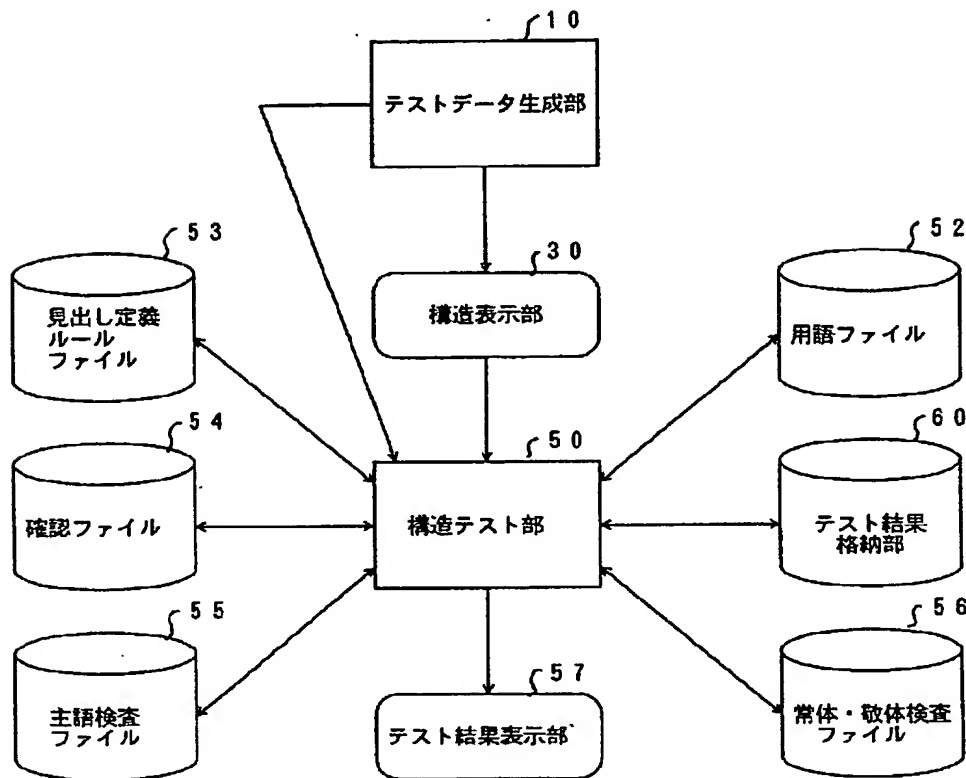
59

分口	和成結果	見出しの部になる用語
項目1	項目1	項目1の部になる用語
項目2	項目2	項目2の部になる用語
項目3	項目3	項目3の部になる用語
項目4	項目4	項目4の部になる用語
項目5	項目5	項目5の部になる用語
項目6	項目6	項目6の部になる用語
項目7	項目7	項目7の部になる用語
項目8	項目8	項目8の部になる用語
項目9	項目9	項目9の部になる用語
項目10	項目10	項目10の部になる用語
項目11	項目11	項目11の部になる用語
項目12	項目12	項目12の部になる用語
項目13	項目13	項目13の部になる用語
項目14	項目14	項目14の部になる用語
項目15	項目15	項目15の部になる用語
項目16	項目16	項目16の部になる用語
項目17	項目17	項目17の部になる用語
項目18	項目18	項目18の部になる用語
項目19	項目19	項目19の部になる用語
項目20	項目20	項目20の部になる用語
項目21	項目21	項目21の部になる用語
項目22	項目22	項目22の部になる用語
項目23	項目23	項目23の部になる用語
項目24	項目24	項目24の部になる用語
項目25	項目25	項目25の部になる用語
項目26	項目26	項目26の部になる用語
項目27	項目27	項目27の部になる用語
項目28	項目28	項目28の部になる用語
項目29	項目29	項目29の部になる用語
項目30	項目30	項目30の部になる用語
項目31	項目31	項目31の部になる用語
項目32	項目32	項目32の部になる用語
項目33	項目33	項目33の部になる用語
項目34	項目34	項目34の部になる用語
項目35	項目35	項目35の部になる用語
項目36	項目36	項目36の部になる用語
項目37	項目37	項目37の部になる用語
項目38	項目38	項目38の部になる用語
項目39	項目39	項目39の部になる用語
項目40	項目40	項目40の部になる用語
項目41	項目41	項目41の部になる用語
項目42	項目42	項目42の部になる用語
項目43	項目43	項目43の部になる用語
項目44	項目44	項目44の部になる用語
項目45	項目45	項目45の部になる用語
項目46	項目46	項目46の部になる用語
項目47	項目47	項目47の部になる用語
項目48	項目48	項目48の部になる用語
項目49	項目49	項目49の部になる用語
項目50	項目50	項目50の部になる用語
項目51	項目51	項目51の部になる用語
項目52	項目52	項目52の部になる用語
項目53	項目53	項目53の部になる用語
項目54	項目54	項目54の部になる用語
項目55	項目55	項目55の部になる用語
項目56	項目56	項目56の部になる用語
項目57	項目57	項目57の部になる用語
項目58	項目58	項目58の部になる用語
項目59	項目59	項目59の部になる用語
項目60	項目60	項目60の部になる用語
項目61	項目61	項目61の部になる用語
項目62	項目62	項目62の部になる用語
項目63	項目63	項目63の部になる用語
項目64	項目64	項目64の部になる用語
項目65	項目65	項目65の部になる用語
項目66	項目66	項目66の部になる用語
項目67	項目67	項目67の部になる用語
項目68	項目68	項目68の部になる用語
項目69	項目69	項目69の部になる用語
項目70	項目70	項目70の部になる用語
項目71	項目71	項目71の部になる用語
項目72	項目72	項目72の部になる用語
項目73	項目73	項目73の部になる用語
項目74	項目74	項目74の部になる用語
項目75	項目75	項目75の部になる用語
項目76	項目76	項目76の部になる用語
項目77	項目77	項目77の部になる用語
項目78	項目78	項目78の部になる用語
項目79	項目79	項目79の部になる用語
項目80	項目80	項目80の部になる用語
項目81	項目81	項目81の部になる用語
項目82	項目82	項目82の部になる用語
項目83	項目83	項目83の部になる用語
項目84	項目84	項目84の部になる用語
項目85	項目85	項目85の部になる用語
項目86	項目86	項目86の部になる用語
項目87	項目87	項目87の部になる用語
項目88	項目88	項目88の部になる用語
項目89	項目89	項目89の部になる用語
項目90	項目90	項目90の部になる用語
項目91	項目91	項目91の部になる用語
項目92	項目92	項目92の部になる用語
項目93	項目93	項目93の部になる用語
項目94	項目94	項目94の部になる用語
項目95	項目95	項目95の部になる用語
項目96	項目96	項目96の部になる用語
項目97	項目97	項目97の部になる用語
項目98	項目98	項目98の部になる用語
項目99	項目99	項目99の部になる用語
項目100	項目100	項目100の部になる用語



【図 7】

本発明の一実施例の構造テスト部の動作を説明するための図



【図 17】

本発明の一実施例のテスト結果を示す図

(A)

【誤り箇所指定】			個数
見出し			1
<input type="checkbox"/> ふさわしくない見出し			1
<input type="checkbox"/> 概念位置の乱れ			0
挿入部			0
<input type="checkbox"/> 書き出しの相違			0
<input type="checkbox"/> 文末の乱れ			0

(B)

ページ	行	内容	個数
P 5	L 2	ふさわしくない見出し	1
		1. 2. 5 である	
ページ	行	概念位置の乱れ	0
P 492	L 18	解説各欄なのに章の見出しに「メリット」があります。	0
ページ	行	書き出しの相違	0
P 115	L 50	章の導入部が「ここでは」で始まっていません。	0
P 200	L 9	節の導入部が「この一では」で始まっていません。	0
ページ	行	文末の乱れ	0
P 1000	L 9	文末が「常体」になっています。	0

【図 18】

本発明の一実施例の修正部の動作を説明するための図

(A)

【誤り箇所指定】			個数
見出し			0
<input type="checkbox"/> 概念位置の乱れ			0
挿入部			0
<input type="checkbox"/> 書き出しの相違			0
<input type="checkbox"/> 文末の乱れ			0

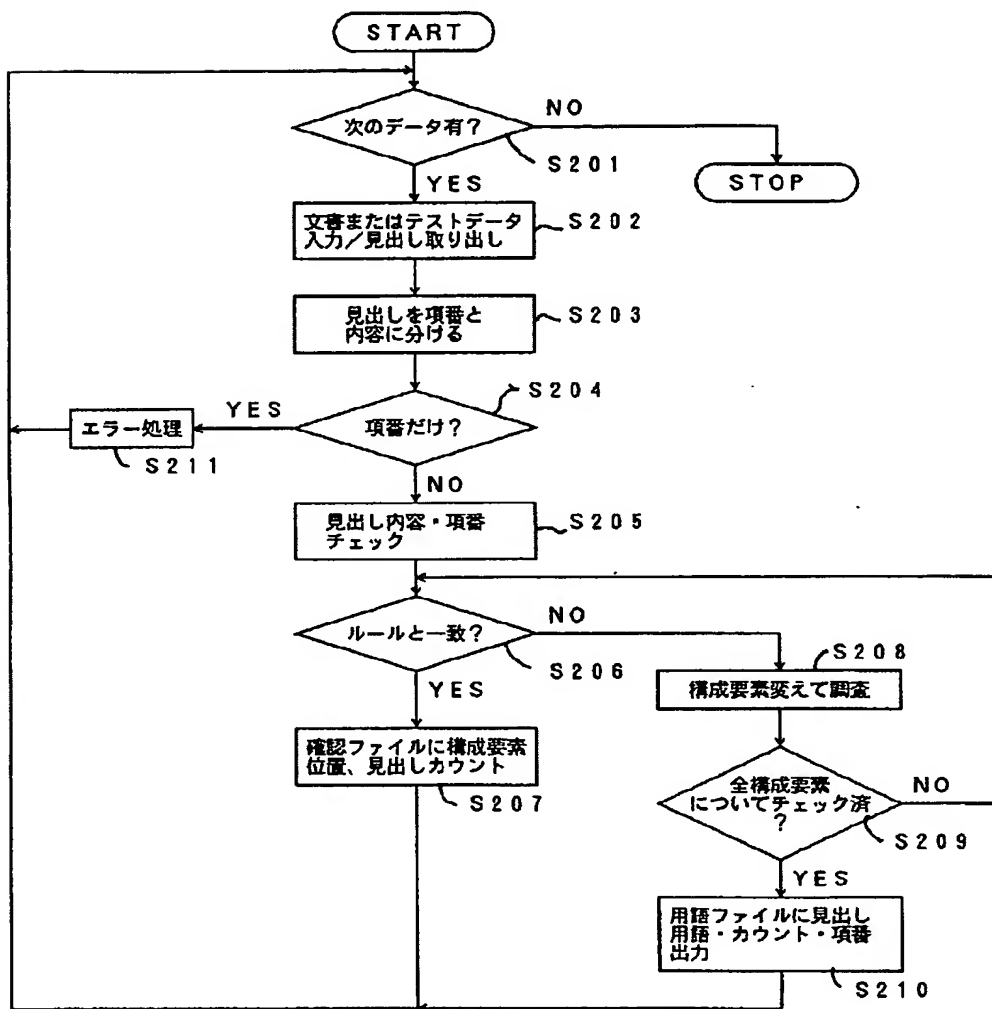
修正したい箇所の表示を指定する。

(B)

ページ	行	内容	個数
P 492	L 18	概念位置の乱れ	0
		第3章 TDS/MAPLEのメリット	
ページ	行	書き出しの相違	0
P 115	L 50	第2では、.....	0
ページ	行	文末の乱れ	0
P 1000	L 9であった。	0

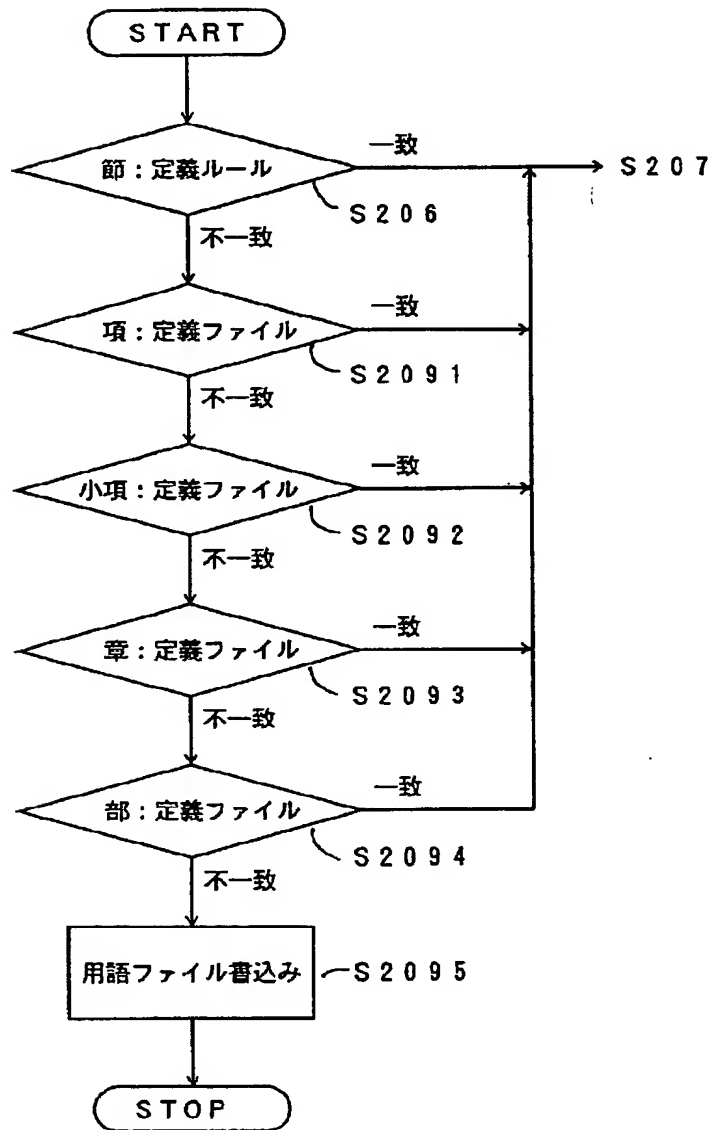
【図9】

本発明の一実施例の構造テスト部の見出しの
テストの概要の動作を示すフローチャート



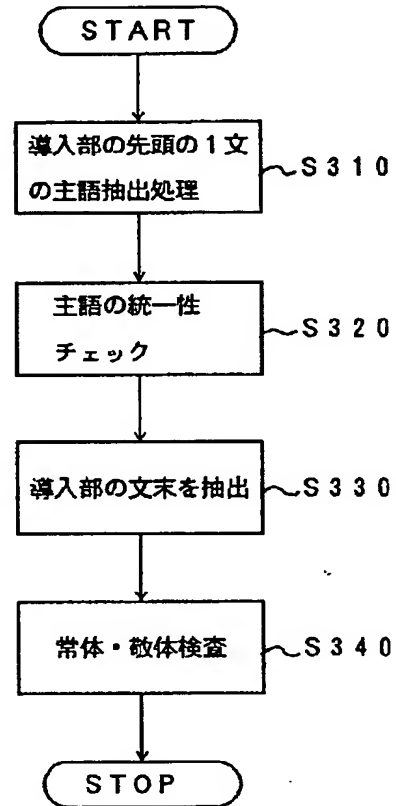
【図10】

本発明の一実施例の用語ファイルに構成要素毎の単語を出力する動作のフローチャート



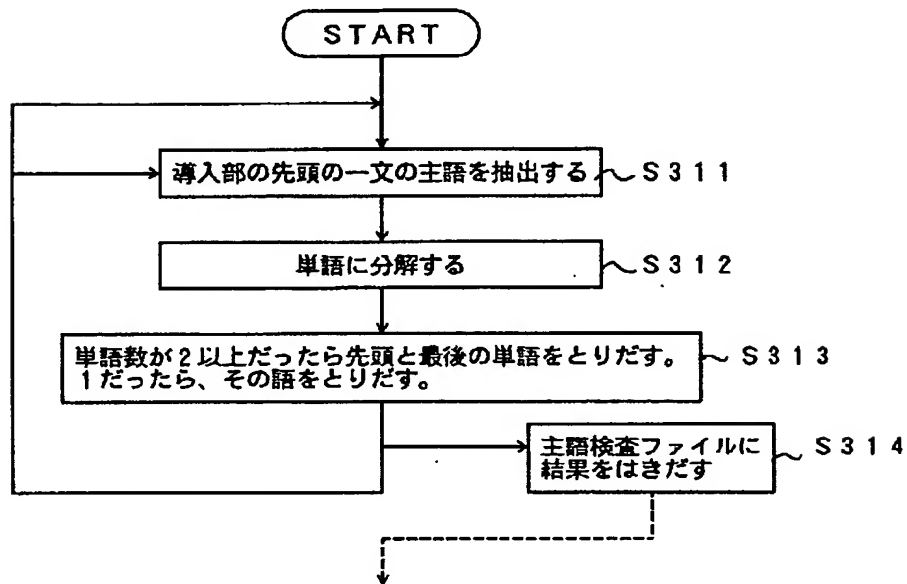
【図11】

本発明の一実施例の構造テスト部の導入部のテストの概要動作を示すフローチャート



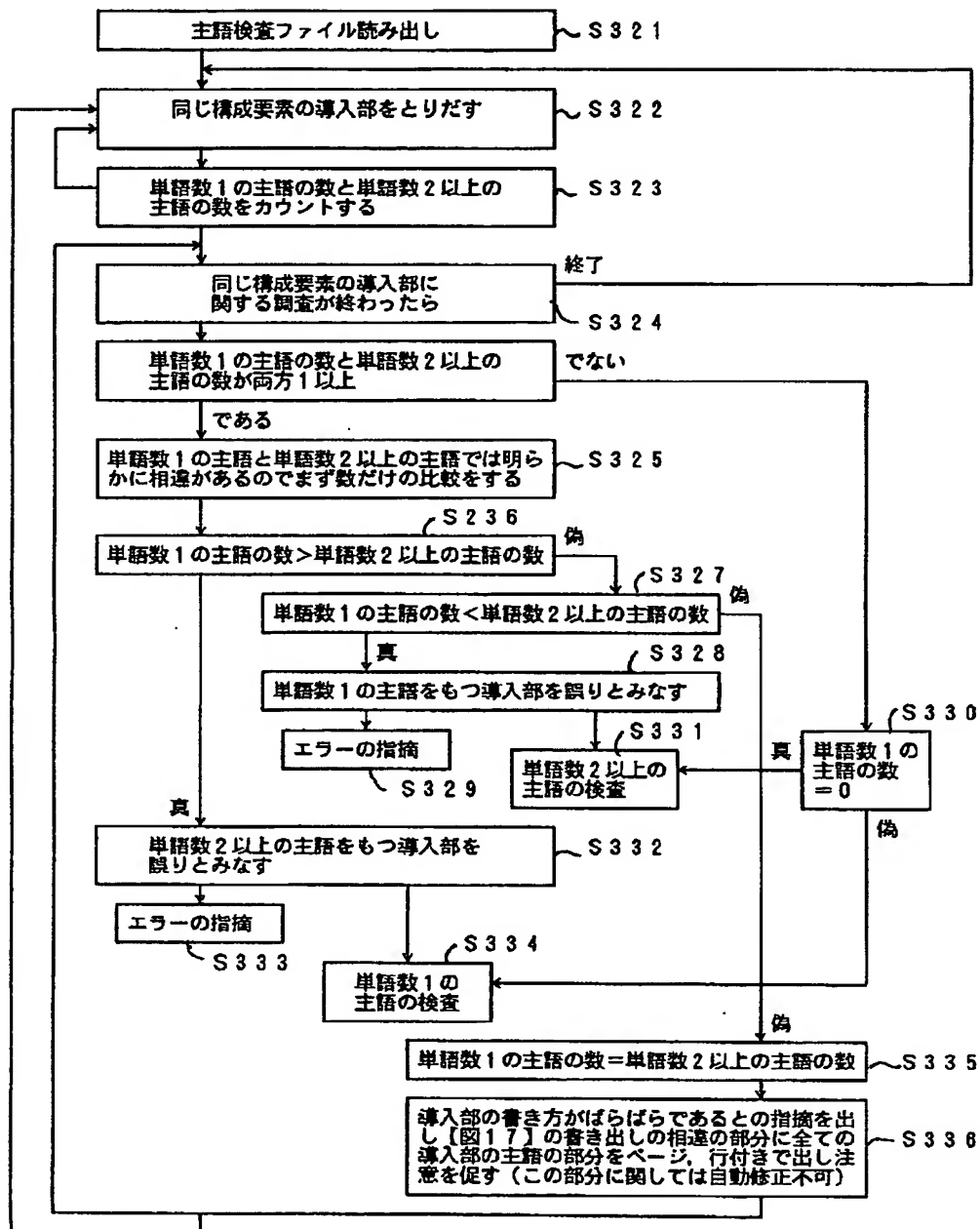
【図 12】

本発明の一実施例の構造テスト部の導入部のテストの
動作を示すフローチャート（その1）



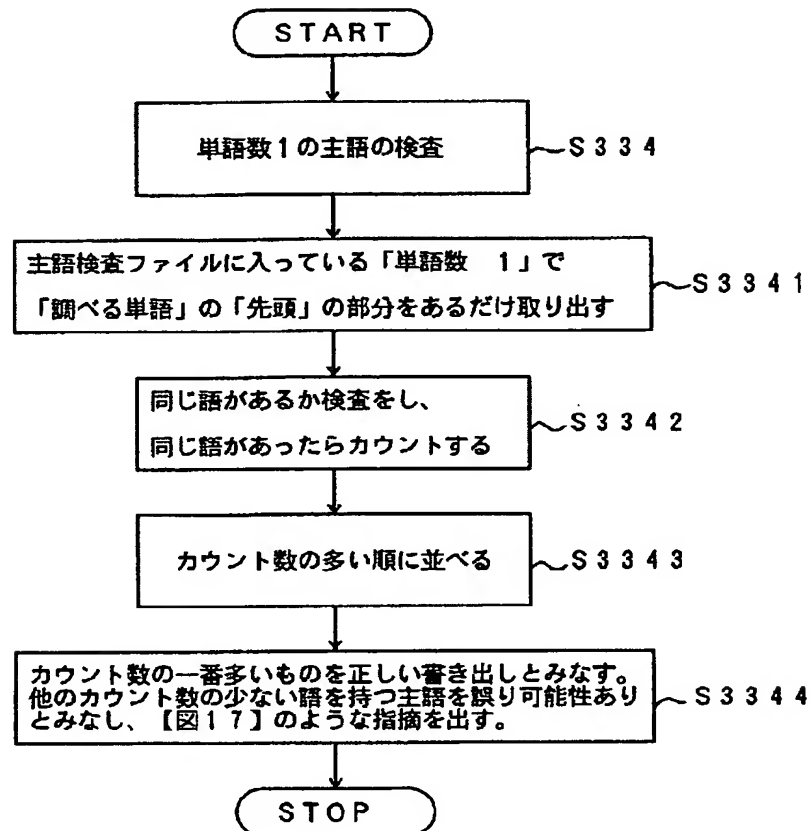
【図13】

本発明の一実施例の構造テスト部の導入部のテストの動作を示すフローチャート（その2）



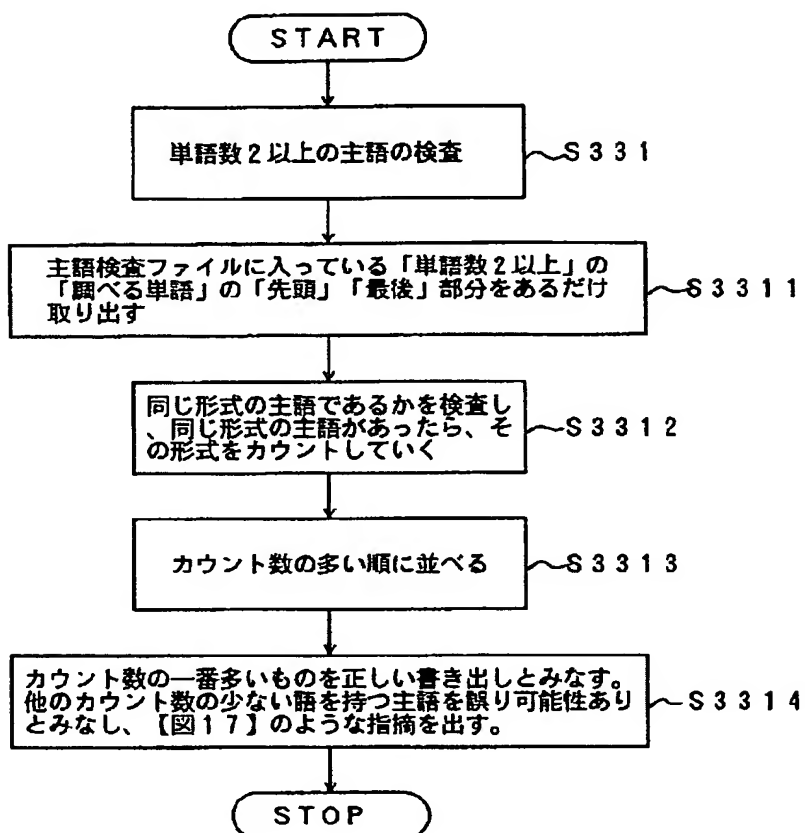
【図 14】

本発明の一実施例の単語数 1 の主語の検査方法を示すフローチャート



【図15】

本発明の一実施例の単語数2以上の主語の検査方法を示すフローチャート



【図 16】

本発明の一実施例の導入部の文末のテストの動作を示すフローチャート

